Linear Algebra

and its applications

FOURTH EDITION

David C. Lay

2. Row reduce the system's augmented matrix:

$$\begin{bmatrix} 1 & -2 & -1 & 3 & 0 \\ -2 & 4 & 5 & -5 & 3 \\ 3 & -6 & -6 & 8 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & -1 & 3 & 0 \\ 0 & 0 & 3 & 1 & 3 \\ 0 & 0 & -3 & -1 & 2 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & -2 & -1 & 3 & 0 \\ 0 & 0 & 3 & 1 & 3 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix}$$

This echelon matrix shows that the system is *inconsistent*, because its rightmost column is a pivot column; the third row corresponds to the equation $0 = 5$. There is no need to perform any more row operations. Note that the presence of the free variables in this problem is irrelevant because the system is inconsistent.

1.3 VECTOR EQUATIONS

Important properties of linear systems can be described with the concept and notation of vectors. This section connects equations involving vectors to ordinary systems of equations. The term *vector* appears in a variety of mathematical and physical contexts, which we will discuss in Chapter 4, "Vector Spaces." Until then, *vector* will mean an *ordered list of numbers*. This simple idea enables us to get to interesting and important applications as quickly as possible.

Vectors in $\mathbb{R}^2$

A matrix with only one column is called a **column vector**, or simply a **vector**. Examples of vectors with two entries are

$$\mathbf{u} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}, \qquad \mathbf{v} = \begin{bmatrix} .2 \\ .3 \end{bmatrix}, \qquad \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

where w_1 and w_2 are any real numbers. The set of all vectors with two entries is denoted by $\mathbb{R}^2$ (read "r-two"). The $\mathbb{R}$ stands for the real numbers that appear as entries in the vectors, and the exponent 2 indicates that each vector contains two entries.[1]

Two vectors in $\mathbb{R}^2$ are **equal** if and only if their corresponding entries are equal. Thus $\begin{bmatrix} 4 \\ 7 \end{bmatrix}$ and $\begin{bmatrix} 7 \\ 4 \end{bmatrix}$ are *not* equal, because vectors in $\mathbb{R}^2$ are *ordered pairs* of real numbers.

Given two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^2$, their **sum** is the vector $\mathbf{u} + \mathbf{v}$ obtained by adding corresponding entries of $\mathbf{u}$ and $\mathbf{v}$. For example,

$$\begin{bmatrix} 1 \\ -2 \end{bmatrix} + \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 1+2 \\ -2+5 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$$

Given a vector $\mathbf{u}$ and a real number c, the **scalar multiple** of $\mathbf{u}$ by c is the vector $c\mathbf{u}$ obtained by multiplying each entry in $\mathbf{u}$ by c. For instance,

$$\text{if} \quad \mathbf{u} = \begin{bmatrix} 3 \\ -1 \end{bmatrix} \quad \text{and} \quad c = 5, \qquad \text{then} \quad c\mathbf{u} = 5\begin{bmatrix} 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 15 \\ -5 \end{bmatrix}$$

[1] Most of the text concerns vectors and matrices that have only real entries. However, all definitions and theorems in Chapters 1–5, and in most of the rest of the text, remain valid if the entries are complex numbers. Complex vectors and matrices arise naturally, for example, in electrical engineering and physics.

The number c in $c\mathbf{u}$ is called a **scalar**; it is written in lightface type to distinguish it from the boldface vector $\mathbf{u}$.

The operations of scalar multiplication and vector addition can be combined, as in the following example.

EXAMPLE 1 Given $\mathbf{u} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 2 \\ -5 \end{bmatrix}$, find $4\mathbf{u}$, $(-3)\mathbf{v}$, and $4\mathbf{u} + (-3)\mathbf{v}$.

SOLUTION

$$4\mathbf{u} = \begin{bmatrix} 4 \\ -8 \end{bmatrix}, \qquad (-3)\mathbf{v} = \begin{bmatrix} -6 \\ 15 \end{bmatrix}$$

and

$$4\mathbf{u} + (-3)\mathbf{v} = \begin{bmatrix} 4 \\ -8 \end{bmatrix} + \begin{bmatrix} -6 \\ 15 \end{bmatrix} = \begin{bmatrix} -2 \\ 7 \end{bmatrix}$$

Sometimes, for convenience (and also to save space), this text may write a column vector such as $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$ in the form $(3, -1)$. In this case, the parentheses and the comma distinguish the vector $(3, -1)$ from the 1×2 row matrix $\begin{bmatrix} 3 & -1 \end{bmatrix}$, written with brackets and no comma. Thus

$$\begin{bmatrix} 3 \\ -1 \end{bmatrix} \neq \begin{bmatrix} 3 & -1 \end{bmatrix}$$

because the matrices have different shapes, even though they have the same entries.

Geometric Descriptions of $\mathbb{R}^2$

Consider a rectangular coordinate system in the plane. Because each point in the plane is determined by an ordered pair of numbers, *we can identify a geometric point* (a, b) *with the column vector* $\begin{bmatrix} a \\ b \end{bmatrix}$. So we may regard $\mathbb{R}^2$ as the set of all points in the plane. See Fig. 1.

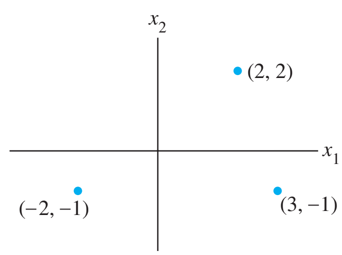


FIGURE 1 Vectors as points.

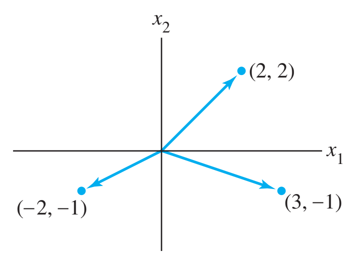


FIGURE 2 Vectors with arrows.

The geometric visualization of a vector such as $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$ is often aided by including an arrow (directed line segment) from the origin $(0, 0)$ to the point $(3, -1)$, as in Fig. 2. In this case, the individual points along the arrow itself have no special significance.[2]

The sum of two vectors has a useful geometric representation. The following rule can be verified by analytic geometry.

[2] In physics, arrows can represent forces and usually are free to move about in space. This interpretation of vectors will be discussed in Section 4.1.

Parallelogram Rule for Addition

If $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^2$ are represented as points in the plane, then $\mathbf{u}+\mathbf{v}$ corresponds to the fourth vertex of the parallelogram whose other vertices are $\mathbf{u}$, $\mathbf{0}$, and $\mathbf{v}$. See Fig. 3.

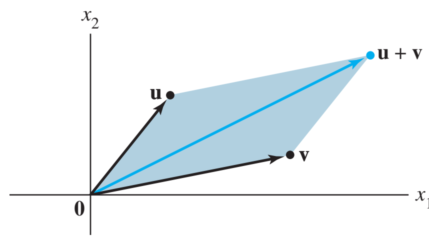


FIGURE 3 The parallelogram rule.

EXAMPLE 2 The vectors $\mathbf{u}=\begin{bmatrix}2\\2\end{bmatrix}$, $\mathbf{v}=\begin{bmatrix}-6\\1\end{bmatrix}$, and $\mathbf{u}+\mathbf{v}=\begin{bmatrix}-4\\3\end{bmatrix}$ are displayed in Fig. 4. ■

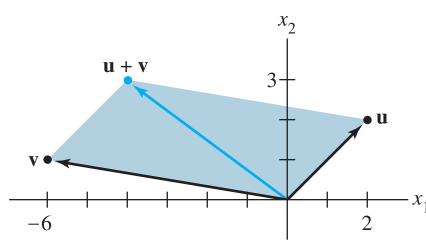
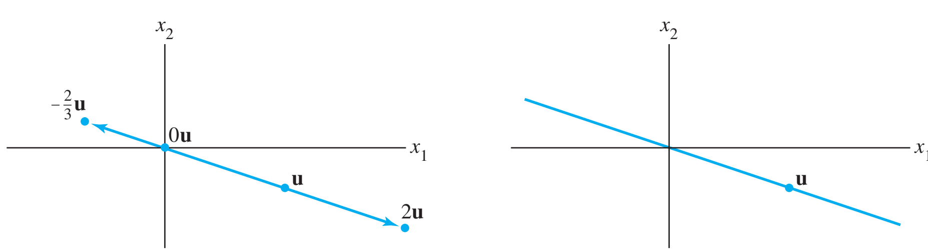


FIGURE 4

The next example illustrates the fact that the set of all scalar multiples of one fixed nonzero vector is a line through the origin, (0, 0).

EXAMPLE 3 Let $\mathbf{u}=\begin{bmatrix}3\\-1\end{bmatrix}$. Display the vectors $\mathbf{u}$, $2\mathbf{u}$, and $-\frac{2}{3}\mathbf{u}$ on a graph.

SOLUTION See Fig. 5, where $\mathbf{u}$, $2\mathbf{u}=\begin{bmatrix}6\\-2\end{bmatrix}$, and $-\frac{2}{3}\mathbf{u}=\begin{bmatrix}-2\\2/3\end{bmatrix}$ are displayed. The arrow for $2\mathbf{u}$ is twice as long as the arrow for $\mathbf{u}$, and the arrows point in the same direction. The arrow for $-\frac{2}{3}\mathbf{u}$ is two-thirds the length of the arrow for $\mathbf{u}$, and the arrows point in opposite directions. In general, the length of the arrow for $c\mathbf{u}$ is $|c|$ times the



Typical multiples of **u** The set of all multiples of **u**

FIGURE 5

length of the arrow for $\mathbf{u}$. [Recall that the length of the line segment from $(0, 0)$ to (a, b) is $\sqrt{a^2 + b^2}$. We shall discuss this further in Chapter 6.] ■

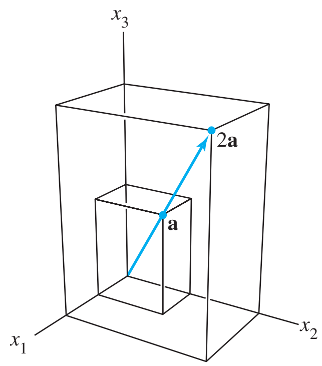


FIGURE 6
Scalar multiples .

Vectors in $\mathbb{R}^3$

Vectors in $\mathbb{R}^3$ are 3×1 column matrices with three entries. They are represented geometrically by points in a three-dimensional coordinate space, with arrows from the origin sometimes included for visual clarity. The vectors $\mathbf{a} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$ and $2\mathbf{a}$ are displayed in Fig. 6.

Vectors in $\mathbb{R}^n$

If n is a positive integer, $\mathbb{R}^n$ (read "r-n") denotes the collection of all lists (or *ordered n-tuples*) of n real numbers, usually written as $n \times 1$ column matrices, such as

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

The vector whose entries are all zero is called the **zero vector** and is denoted by $\mathbf{0}$. (The number of entries in $\mathbf{0}$ will be clear from the context.)

Equality of vectors in $\mathbb{R}^n$ and the operations of scalar multiplication and vector addition in $\mathbb{R}^n$ are defined entry by entry just as in $\mathbb{R}^2$. These operations on vectors have the following properties, which can be verified directly from the corresponding properties for real numbers. See Practice Problem 1 and Exercises 33 and 34 at the end of this section.

Algebraic Properties of $\mathbb{R}^n$

For all $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in $\mathbb{R}^n$ and all scalars c and d:

(i) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
(ii) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
(iii) $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$
(iv) $\mathbf{u} + (-\mathbf{u}) = -\mathbf{u} + \mathbf{u} = \mathbf{0}$, where $-\mathbf{u}$ denotes $(-1)\mathbf{u}$
(v) $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$
(vi) $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$
(vii) $c(d\mathbf{u}) = (cd)(\mathbf{u})$
(viii) $1\mathbf{u} = \mathbf{u}$

For simplicity of notation, a vector such as $\mathbf{u} + (-1)\mathbf{v}$ is often written as $\mathbf{u} - \mathbf{v}$. Figure 7 shows $\mathbf{u} - \mathbf{v}$ as the sum of $\mathbf{u}$ and $-\mathbf{v}$.

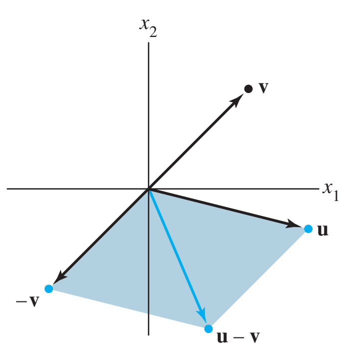


FIGURE 7
Vector subtraction.

Linear Combinations

Given vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p$ in $\mathbb{R}^n$ and given scalars $c_1, c_2, \dots, c_p$, the vector $\mathbf{y}$ defined by

$$\mathbf{y} = c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p$$

is called a **linear combination** of $\mathbf{v}_1, \dots, \mathbf{v}_p$ with **weights** $c_1, \dots, c_p$. Property (ii) above permits us to omit parentheses when forming such a linear combination. The weights in a linear combination can be any real numbers, including zero. For example, some linear combinations of vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are

$$\sqrt{3}\,\mathbf{v}_1 + \mathbf{v}_2, \quad \tfrac{1}{2}\mathbf{v}_1 \;(= \tfrac{1}{2}\mathbf{v}_1 + 0\mathbf{v}_2), \quad \text{and} \quad \mathbf{0}\;(= 0\mathbf{v}_1 + 0\mathbf{v}_2)$$

EXAMPLE 4 Figure 8 identifies selected linear combinations of $\mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. (Note that sets of parallel grid lines are drawn through integer multiples of $\mathbf{v}_1$ and $\mathbf{v}_2$.) Estimate the linear combinations of $\mathbf{v}_1$ and $\mathbf{v}_2$ that generate the vectors $\mathbf{u}$ and $\mathbf{w}$.

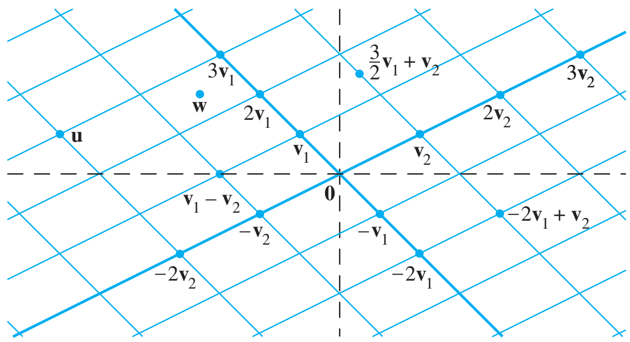


FIGURE 8 Linear combinations of $\mathbf{v}_1$ and $\mathbf{v}_2$.

SOLUTION The parallelogram rule shows that $\mathbf{u}$ is the sum of $3\mathbf{v}_1$ and $-2\mathbf{v}_2$; that is,

$$\mathbf{u} = 3\mathbf{v}_1 - 2\mathbf{v}_2$$

This expression for $\mathbf{u}$ can be interpreted as instructions for traveling from the origin to $\mathbf{u}$ along two straight paths. First, travel 3 units in the $\mathbf{v}_1$ direction to $3\mathbf{v}_1$, and then travel -2 units in the $\mathbf{v}_2$ direction (parallel to the line through $\mathbf{v}_2$ and $\mathbf{0}$). Next, although the vector $\mathbf{w}$ is not on a grid line, $\mathbf{w}$ appears to be about halfway between two pairs of grid lines, at the vertex of a parallelogram determined by $(5/2)\mathbf{v}_1$ and $(-1/2)\mathbf{v}_2$. (See Fig. 9.) Thus a reasonable estimate for $\mathbf{w}$ is

$$\mathbf{w} = \tfrac{5}{2}\mathbf{v}_1 - \tfrac{1}{2}\mathbf{v}_2$$

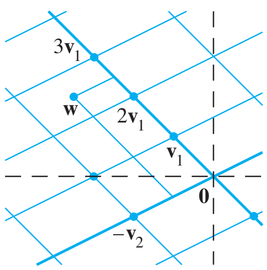


FIGURE 9

The next example connects a problem about linear combinations to the fundamental existence question studied in Sections 1.1 and 1.2.

EXAMPLE 5 Let $\mathbf{a}_1 = \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix}$, and $\mathbf{b} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$. Determine whether $\mathbf{b}$ can be generated (or written) as a linear combination of $\mathbf{a}_1$ and $\mathbf{a}_2$. That is, determine whether weights x_1 and x_2 exist such that

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 = \mathbf{b} \tag{1}$$

If vector equation (1) has a solution, find it.

SOLUTION Use the definitions of scalar multiplication and vector addition to rewrite the vector equation

$$x_1 \underset{\mathbf{a}_1}{\underset{\uparrow}{\begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix}}} + x_2 \underset{\mathbf{a}_2}{\underset{\uparrow}{\begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix}}} = \underset{\mathbf{b}}{\underset{\uparrow}{\begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}}}$$

which is the same as

$$\begin{bmatrix} x_1 \\ -2x_1 \\ -5x_1 \end{bmatrix} + \begin{bmatrix} 2x_2 \\ 5x_2 \\ 6x_2 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$$

and

$$\begin{bmatrix} x_1 + 2x_2 \\ -2x_1 + 5x_2 \\ -5x_1 + 6x_2 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix} \tag{2}$$

The vectors on the left and right sides of (2) are equal if and only if their corresponding entries are both equal. That is, x_1 and x_2 make the vector equation (1) true if and only if x_1 and x_2 satisfy the system

$$\begin{aligned} x_1 + 2x_2 &= 7 \\ -2x_1 + 5x_2 &= 4 \\ -5x_1 + 6x_2 &= -3 \end{aligned} \tag{3}$$

To solve this system, row reduce the augmented matrix of the system as follows:[3]

$$\begin{bmatrix} 1 & 2 & 7 \\ -2 & 5 & 4 \\ -5 & 6 & -3 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 7 \\ 0 & 9 & 18 \\ 0 & 16 & 32 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 7 \\ 0 & 1 & 2 \\ 0 & 16 & 32 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

The solution of (3) is $x_1 = 3$ and $x_2 = 2$. Hence **b** is a linear combination of $\mathbf{a}_1$ and $\mathbf{a}_2$, with weights $x_1 = 3$ and $x_2 = 2$. That is,

$$3\begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix} + 2\begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$$

Observe in Example 5 that the original vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and **b** are the columns of the augmented matrix that we row reduced:

$$\begin{array}{ccc} \begin{bmatrix} 1 & 2 & 7 \\ -2 & 5 & 4 \\ -5 & 6 & -3 \end{bmatrix} \\ \uparrow \quad \uparrow \quad \uparrow \\ \mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{b} \end{array}$$

For brevity, write this matrix in a way that identifies its columns—namely,

$$\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{b} \end{bmatrix} \tag{4}$$

It is clear how to write this augmented matrix immediately from vector equation (1), without going through the intermediate steps of Example 5. Take the vectors in the order in which they appear in (1) and put them into the columns of a matrix as in (4).

The discussion above is easily modified to establish the following fundamental fact.

A vector equation

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$$

has the same solution set as the linear system whose augmented matrix is

$$\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n & \mathbf{b} \end{bmatrix} \tag{5}$$

In particular, **b** can be generated by a linear combination of $\mathbf{a}_1, \dots, \mathbf{a}_n$ if and only if there exists a solution to the linear system corresponding to the matrix (5).

[3] The symbol $\sim$ between matrices denotes row equivalence (Section 1.2).

One of the key ideas in linear algebra is to study the set of all vectors that can be generated or written as a linear combination of a fixed set $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ of vectors.

DEFINITION

If $\mathbf{v}_1, \dots, \mathbf{v}_p$ are in $\mathbb{R}^n$, then the set of all linear combinations of $\mathbf{v}_1, \dots, \mathbf{v}_p$ is denoted by Span $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ and is called the **subset of $\mathbb{R}^n$ spanned** (or **generated**) **by** $\mathbf{v}_1, \dots, \mathbf{v}_p$. That is, Span $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ is the collection of all vectors that can be written in the form

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p$$

with $c_1, \dots, c_p$ scalars.

Asking whether a vector $\mathbf{b}$ is in Span $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ amounts to asking whether the vector equation

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p = \mathbf{b}$$

has a solution, or, equivalently, asking whether the linear system with augmented matrix $[\,\mathbf{v}_1 \;\cdots\; \mathbf{v}_p \;\; \mathbf{b}\,]$ has a solution.

Note that Span $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ contains every scalar multiple of $\mathbf{v}_1$ (for example), since $c\mathbf{v}_1 = c\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_p$. In particular, the zero vector must be in Span $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$.

A Geometric Description of Span $\{\mathbf{v}\}$ and Span $\{\mathbf{u}, \mathbf{v}\}$

Let $\mathbf{v}$ be a nonzero vector in $\mathbb{R}^3$. Then Span $\{\mathbf{v}\}$ is the set of all scalar multiples of $\mathbf{v}$, which is the set of points on the line in $\mathbb{R}^3$ through $\mathbf{v}$ and $\mathbf{0}$. See Fig. 10.

If $\mathbf{u}$ and $\mathbf{v}$ are nonzero vectors in $\mathbb{R}^3$, with $\mathbf{v}$ not a multiple of $\mathbf{u}$, then Span $\{\mathbf{u}, \mathbf{v}\}$ is the plane in $\mathbb{R}^3$ that contains $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{0}$. In particular, Span $\{\mathbf{u}, \mathbf{v}\}$ contains the line in $\mathbb{R}^3$ through $\mathbf{u}$ and $\mathbf{0}$ and the line through $\mathbf{v}$ and $\mathbf{0}$. See Fig. 11.

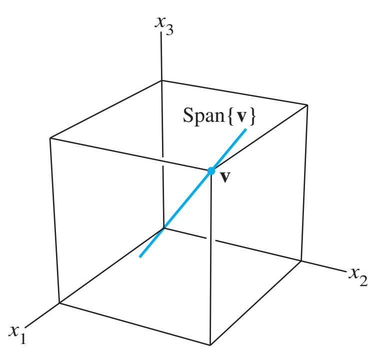


FIGURE 10 Span $\{\mathbf{v}\}$ as a line through the origin.

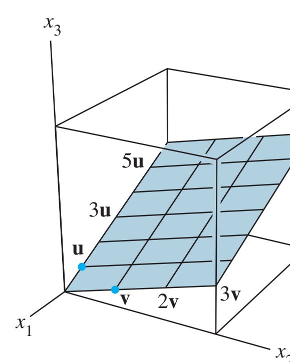


FIGURE 11 Span $\{\mathbf{u}, \mathbf{v}\}$ as a plane through the origin.

EXAMPLE 6 Let $\mathbf{a}_1 = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix} 5 \\ -13 \\ -3 \end{bmatrix}$, and $\mathbf{b} = \begin{bmatrix} -3 \\ 8 \\ 1 \end{bmatrix}$. Then Span $\{\mathbf{a}_1, \mathbf{a}_2\}$ is a plane through the origin in $\mathbb{R}^3$. Is $\mathbf{b}$ in that plane?

SOLUTION Does the equation $x_1\mathbf{a}_1 + x_2\mathbf{a}_2 = \mathbf{b}$ have a solution? To answer this, row reduce the augmented matrix $[\,\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{b}\,]$:

$$\begin{bmatrix} 1 & 5 & -3 \\ -2 & -13 & 8 \\ 3 & -3 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & -3 \\ 0 & -3 & 2 \\ 0 & -18 & 10 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & -3 \\ 0 & -3 & 2 \\ 0 & 0 & -2 \end{bmatrix}$$

The third equation is $0 = -2$, which shows that the system has no solution. The vector equation $x_1\mathbf{a}_1 + x_2\mathbf{a}_2 = \mathbf{b}$ has no solution, and so $\mathbf{b}$ is *not* in Span $\{\mathbf{a}_1, \mathbf{a}_2\}$. ■

Linear Combinations in Applications

The final example shows how scalar multiples and linear combinations can arise when a quantity such as "cost" is broken down into several categories. The basic principle for the example concerns the cost of producing several units of an item when the cost per unit is known:

$$\left\{\begin{matrix}\text{number} \\ \text{of units}\end{matrix}\right\} \cdot \left\{\begin{matrix}\text{cost} \\ \text{per unit}\end{matrix}\right\} = \left\{\begin{matrix}\text{total} \\ \text{cost}\end{matrix}\right\}$$

EXAMPLE 7 A company manufactures two products. For \$1.00 worth of product B, the company spends \$.45 on materials, \$.25 on labor, and \$.15 on overhead. For \$1.00 worth of product C, the company spends \$.40 on materials, \$.30 on labor, and \$.15 on overhead. Let

$$\mathbf{b} = \begin{bmatrix} .45 \\ .25 \\ .15 \end{bmatrix} \quad \text{and} \quad \mathbf{c} = \begin{bmatrix} .40 \\ .30 \\ .15 \end{bmatrix}$$

Then $\mathbf{b}$ and $\mathbf{c}$ represent the "costs per dollar of income" for the two products.

a. What economic interpretation can be given to the vector $100\mathbf{b}$?

b. Suppose the company wishes to manufacture x_1 dollars worth of product B and x_2 dollars worth of product C. Give a vector that describes the various costs the company will have (for materials, labor, and overhead).

SOLUTION

a. Compute

$$100\mathbf{b} = 100\begin{bmatrix} .45 \\ .25 \\ .15 \end{bmatrix} = \begin{bmatrix} 45 \\ 25 \\ 15 \end{bmatrix}$$

The vector $100\mathbf{b}$ lists the various costs for producing \$100 worth of product B—namely, \$45 for materials, \$25 for labor, and \$15 for overhead.

b. The costs of manufacturing x_1 dollars worth of B are given by the vector $x_1\mathbf{b}$, and the costs of manufacturing x_2 dollars worth of C are given by $x_2\mathbf{c}$. Hence the total costs for both products are given by the vector $x_1\mathbf{b} + x_2\mathbf{c}$. ■

PRACTICE PROBLEMS

1. Prove that $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ for any $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$.
2. For what value(s) of h will $\mathbf{y}$ be in Span$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ if

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 5 \\ -4 \\ -7 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix}, \quad \text{and} \quad \mathbf{y} = \begin{bmatrix} -4 \\ 3 \\ h \end{bmatrix}$$

1.3 EXERCISES

In Exercises 1 and 2, compute $\mathbf{u}+\mathbf{v}$ and $\mathbf{u}-2\mathbf{v}$.

1. $\mathbf{u}=\begin{bmatrix}-1\\2\end{bmatrix}, \mathbf{v}=\begin{bmatrix}-3\\-1\end{bmatrix}$

2. $\mathbf{u}=\begin{bmatrix}3\\2\end{bmatrix}, \mathbf{v}=\begin{bmatrix}2\\-1\end{bmatrix}$

In Exercises 3 and 4, display the following vectors using arrows on an xy-graph: $\mathbf{u}$, $\mathbf{v}$, $-\mathbf{v}$, $-2\mathbf{v}$, $\mathbf{u}+\mathbf{v}$, $\mathbf{u}-\mathbf{v}$, and $\mathbf{u}-2\mathbf{v}$. Notice that $\mathbf{u}-\mathbf{v}$ is the vertex of a parallelogram whose other vertices are $\mathbf{u}$, $\mathbf{0}$, and $-\mathbf{v}$.

3. $\mathbf{u}$ and $\mathbf{v}$ as in Exercise 1

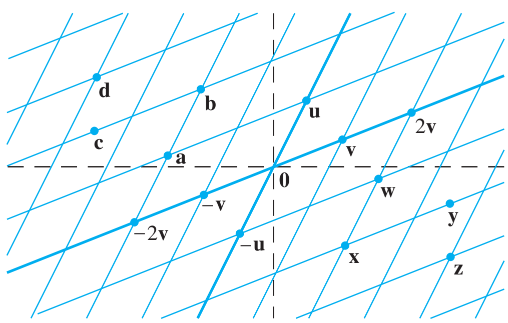
4. $\mathbf{u}$ and $\mathbf{v}$ as in Exercise 2

In Exercises 5 and 6, write a system of equations that is equivalent to the given vector equation.

5. $x_1\begin{bmatrix}3\\-2\\8\end{bmatrix}+x_2\begin{bmatrix}5\\0\\-9\end{bmatrix}=\begin{bmatrix}2\\-3\\8\end{bmatrix}$

6. $x_1\begin{bmatrix}3\\-2\end{bmatrix}+x_2\begin{bmatrix}7\\3\end{bmatrix}+x_3\begin{bmatrix}-2\\1\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$

Use the accompanying figure to write each vector listed in Exercises 7 and 8 as a linear combination of $\mathbf{u}$ and $\mathbf{v}$. Is every vector in $\mathbb{R}^2$ a linear combination of $\mathbf{u}$ and $\mathbf{v}$?



7. Vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, and $\mathbf{d}$

8. Vectors $\mathbf{w}$, $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$

In Exercises 9 and 10, write a vector equation that is equivalent to the given system of equations.

9. $$\begin{aligned} x_2+5x_3&=0\\ 4x_1+6x_2-x_3&=0\\ -x_1+3x_2-8x_3&=0 \end{aligned}$$

10. $$\begin{aligned} 3x_1-2x_2+4x_3&=3\\ -2x_1-7x_2+5x_3&=1\\ 5x_1+4x_2-3x_3&=2 \end{aligned}$$

In Exercises 11 and 12, determine if $\mathbf{b}$ is a linear combination of $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$.

11. $\mathbf{a}_1=\begin{bmatrix}1\\-2\\0\end{bmatrix}, \mathbf{a}_2=\begin{bmatrix}0\\1\\2\end{bmatrix}, \mathbf{a}_3=\begin{bmatrix}5\\-6\\8\end{bmatrix}, \mathbf{b}=\begin{bmatrix}2\\-1\\6\end{bmatrix}$

12. $\mathbf{a}_1=\begin{bmatrix}1\\0\\1\end{bmatrix}, \mathbf{a}_2=\begin{bmatrix}-2\\3\\-2\end{bmatrix}, \mathbf{a}_3=\begin{bmatrix}-6\\7\\5\end{bmatrix}, \mathbf{b}=\begin{bmatrix}11\\-5\\9\end{bmatrix}$

In Exercises 13 and 14, determine if $\mathbf{b}$ is a linear combination of the vectors formed from the columns of the matrix A.

13. $A=\begin{bmatrix}1&-4&2\\0&3&5\\-2&8&-4\end{bmatrix}, \mathbf{b}=\begin{bmatrix}3\\-7\\-3\end{bmatrix}$

14. $A=\begin{bmatrix}1&0&5\\-2&1&-6\\0&2&8\end{bmatrix}, \mathbf{b}=\begin{bmatrix}2\\-1\\6\end{bmatrix}$

15. Let $\mathbf{a}_1=\begin{bmatrix}1\\3\\-1\end{bmatrix}, \mathbf{a}_2=\begin{bmatrix}-5\\-8\\2\end{bmatrix}$, and $\mathbf{b}=\begin{bmatrix}3\\-5\\h\end{bmatrix}$. For what value(s) of h is $\mathbf{b}$ in the plane spanned by $\mathbf{a}_1$ and $\mathbf{a}_2$?

16. Let $\mathbf{v}_1=\begin{bmatrix}1\\0\\-2\end{bmatrix}, \mathbf{v}_2=\begin{bmatrix}-2\\1\\7\end{bmatrix}$, and $\mathbf{y}=\begin{bmatrix}h\\-3\\-5\end{bmatrix}$. For what value(s) of h is $\mathbf{y}$ in the plane generated by $\mathbf{v}_1$ and $\mathbf{v}_2$?

In Exercises 17 and 18, list five vectors in Span $\{\mathbf{v}_1, \mathbf{v}_2\}$. For each vector, show the weights on $\mathbf{v}_1$ and $\mathbf{v}_2$ used to generate the vector and list the three entries of the vector. Do not make a sketch.

17. $\mathbf{v}_1=\begin{bmatrix}3\\1\\2\end{bmatrix}, \mathbf{v}_2=\begin{bmatrix}-4\\0\\1\end{bmatrix}$

18. $\mathbf{v}_1=\begin{bmatrix}1\\1\\-2\end{bmatrix}, \mathbf{v}_2=\begin{bmatrix}-2\\3\\0\end{bmatrix}$

19. Give a geometric description of Span $\{\mathbf{v}_1, \mathbf{v}_2\}$ for the vectors $\mathbf{v}_1=\begin{bmatrix}8\\2\\-6\end{bmatrix}$ and $\mathbf{v}_2=\begin{bmatrix}12\\3\\-9\end{bmatrix}$.

20. Give a geometric description of Span $\{\mathbf{v}_1, \mathbf{v}_2\}$ for the vectors in Exercise 18.

21. Let $\mathbf{u}=\begin{bmatrix}2\\-1\end{bmatrix}$ and $\mathbf{v}=\begin{bmatrix}2\\1\end{bmatrix}$. Show that $\begin{bmatrix}h\\k\end{bmatrix}$ is in Span $\{\mathbf{u}, \mathbf{v}\}$ for all h and k.

22. Construct a 3×3 matrix A, with nonzero entries, and a vector $\mathbf{b}$ in $\mathbb{R}^3$ such that $\mathbf{b}$ is *not* in the set spanned by the columns of A.

In Exercises 23 and 24, mark each statement True or False. Justify each answer.

23. a. Another notation for the vector $\begin{bmatrix}-4\\3\end{bmatrix}$ is $[\,-4 \quad 3\,]$.

    b. The points in the plane corresponding to $\begin{bmatrix}-2\\5\end{bmatrix}$ and $\begin{bmatrix}-5\\2\end{bmatrix}$ lie on a line through the origin.

    c. An example of a linear combination of vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ is the vector $\frac{1}{2}\mathbf{v}_1$.

d. The solution set of the linear system whose augmented matrix is $[\,\mathbf{a}_1 \;\; \mathbf{a}_2 \;\; \mathbf{a}_3 \;\; \mathbf{b}\,]$ is the same as the solution set of the equation $x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3 = \mathbf{b}$.

e. The set Span $\{\mathbf{u}, \mathbf{v}\}$ is always visualized as a plane through the origin.

24. a. When $\mathbf{u}$ and $\mathbf{v}$ are nonzero vectors, Span $\{\mathbf{u}, \mathbf{v}\}$ contains only the line through $\mathbf{u}$ and the origin, and the line through $\mathbf{v}$ and the origin.

b. Any list of five real numbers is a vector in $\mathbb{R}^5$.

c. Asking whether the linear system corresponding to an augmented matrix $[\,\mathbf{a}_1 \;\; \mathbf{a}_2 \;\; \mathbf{a}_3 \;\; \mathbf{b}\,]$ has a solution amounts to asking whether $\mathbf{b}$ is in Span $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$.

d. The vector $\mathbf{v}$ results when a vector $\mathbf{u} - \mathbf{v}$ is added to the vector $\mathbf{v}$.

e. The weights $c_1, \dots, c_p$ in a linear combination $c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p$ cannot all be zero.

25. Let $A = \begin{bmatrix} 1 & 0 & -4 \\ 0 & 3 & -2 \\ -2 & 6 & 3 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 4 \\ 1 \\ -4 \end{bmatrix}$. Denote the columns of A by $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$, and let $W = \text{Span}\,\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$.

a. Is $\mathbf{b}$ in $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$? How many vectors are in $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$?

b. Is $\mathbf{b}$ in W? How many vectors are in W?

c. Show that $\mathbf{a}_1$ is in W. [*Hint:* Row operations are unnecessary.]

26. Let $A = \begin{bmatrix} 2 & 0 & 6 \\ -1 & 8 & 5 \\ 1 & -2 & 1 \end{bmatrix}$, let $\mathbf{b} = \begin{bmatrix} 10 \\ 3 \\ 7 \end{bmatrix}$, and let W be the set of all linear combinations of the columns of A.

a. Is $\mathbf{b}$ in W?

b. Show that the second column of A is in W.

27. A mining company has two mines. One day's operation at mine #1 produces ore that contains 30 metric tons of copper and 600 kilograms of silver, while one day's operation at mine #2 produces ore that contains 40 metric tons of copper and 380 kilograms of silver. Let $\mathbf{v}_1 = \begin{bmatrix} 30 \\ 600 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 40 \\ 380 \end{bmatrix}$. Then $\mathbf{v}_1$ and $\mathbf{v}_2$ represent the "output per day" of mine #1 and mine #2, respectively.

a. What physical interpretation can be given to the vector $5\mathbf{v}_1$?

b. Suppose the company operates mine #1 for x_1 days and mine #2 for x_2 days. Write a vector equation whose solution gives the number of days each mine should operate in order to produce 240 tons of copper and 2824 kilograms of silver. Do not solve the equation.

c. **[M]** Solve the equation in (b).

28. A steam plant burns two types of coal: anthracite (A) and bituminous (B). For each ton of A burned, the plant produces 27.6 million Btu of heat, 3100 grams (g) of sulfur dioxide, and 250 g of particulate matter (solid-particle pollutants). For each ton of B burned, the plant produces 30.2 million Btu, 6400 g of sulfur dioxide, and 360 g of particulate matter.

a. How much heat does the steam plant produce when it burns x_1 tons of A and x_2 tons of B?

b. Suppose the output of the steam plant is described by a vector that lists the amounts of heat, sulfur dioxide, and particulate matter. Express this output as a linear combination of two vectors, assuming that the plant burns x_1 tons of A and x_2 tons of B.

c. **[M]** Over a certain time period, the steam plant produced 162 million Btu of heat, 23,610 g of sulfur dioxide, and 1623 g of particulate matter. Determine how many tons of each type of coal the steam plant must have burned. Include a vector equation as part of your solution.

29. Let $\mathbf{v}_1, \dots, \mathbf{v}_k$ be points in $\mathbb{R}^3$ and suppose that for $j = 1, \dots, k$ an object with mass m_j is located at point $\mathbf{v}_j$. Physicists call such objects *point masses*. The total mass of the system of point masses is

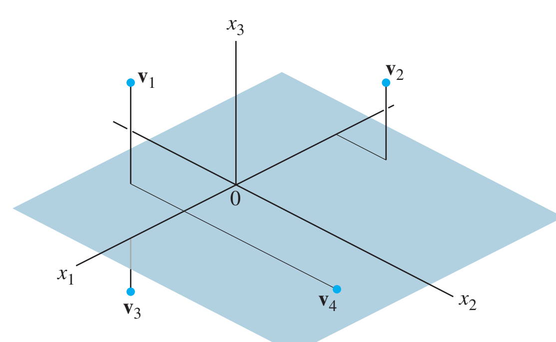
$$m = m_1 + \cdots + m_k$$

The *center of gravity* (or *center of mass*) of the system is

$$\bar{\mathbf{v}} = \frac{1}{m}[m_1\mathbf{v}_1 + \cdots + m_k\mathbf{v}_k]$$

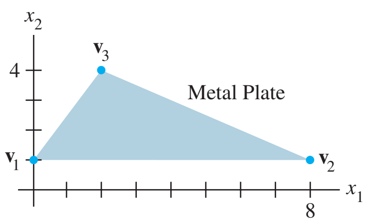
Compute the center of gravity of the system consisting of the following point masses (see the figure):

Point	Mass
$\mathbf{v}_1 = (2, -2, 4)$	4 g
$\mathbf{v}_2 = (-4, 2, 3)$	2 g
$\mathbf{v}_3 = (4, 0, -2)$	3 g
$\mathbf{v}_4 = (1, -6, 0)$	5 g



30. Let $\bar{\mathbf{v}}$ be the center of mass of a system of point masses located at $\mathbf{v}_1, \dots, \mathbf{v}_k$ as in Exercise 29. Is $\bar{\mathbf{v}}$ in Span $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$? Explain.

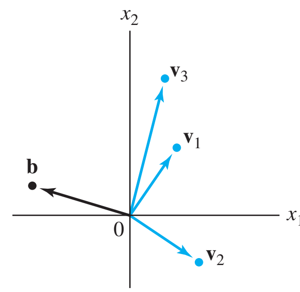
31. A thin triangular plate of uniform density and thickness has vertices at $\mathbf{v}_1 = (0, 1)$, $\mathbf{v}_2 = (8, 1)$, and $\mathbf{v}_3 = (2, 4)$, as in the figure below, and the mass of the plate is 3 g.



a. Find the (x, y)-coordinates of the center of mass of the plate. This "balance point" of the plate coincides with the center of mass of a system consisting of three 1-gram point masses located at the vertices of the plate.

b. Determine how to distribute an additional mass of 6 g at the three vertices of the plate to move the balance point of the plate to $(2, 2)$. [*Hint:* Let w_1, w_2, and w_3 denote the masses added at the three vertices, so that $w_1 + w_2 + w_3 = 6$.]

32. Consider the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, and $\mathbf{b}$ in $\mathbb{R}^2$, shown in the figure. Does the equation $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 = \mathbf{b}$ have a solution? Is the solution unique? Use the figure to explain your answers.



33. Use the vectors $\mathbf{u} = (u_1, \dots, u_n)$, $\mathbf{v} = (v_1, \dots, v_n)$, and $\mathbf{w} = (w_1, \dots, w_n)$ to verify the following algebraic properties of $\mathbb{R}^n$.

a. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$

b. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$ for each scalar c

34. Use the vector $\mathbf{u} = (u_1, \dots, u_n)$ to verify the following algebraic properties of $\mathbb{R}^n$.

a. $\mathbf{u} + (-\mathbf{u}) = (-\mathbf{u}) + \mathbf{u} = \mathbf{0}$

b. $c(d\mathbf{u}) = (cd)\mathbf{u}$ for all scalars c and d

SOLUTIONS TO PRACTICE PROBLEMS

1. Take arbitrary vectors $\mathbf{u} = (u_1, \dots, u_n)$ and $\mathbf{v} = (v_1, \dots, v_n)$ in $\mathbb{R}^n$, and compute

$$\begin{aligned} \mathbf{u} + \mathbf{v} &= (u_1 + v_1, \dots, u_n + v_n) && \text{Definition of vector addition} \\ &= (v_1 + u_1, \dots, v_n + u_n) && \text{Commutativity of addition in } \mathbb{R} \\ &= \mathbf{v} + \mathbf{u} && \text{Definition of vector addition} \end{aligned}$$

2. The vector $\mathbf{y}$ belongs to Span $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ if and only if there exist scalars x_1, x_2, x_3 such that

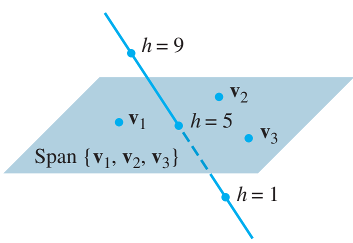
$$x_1 \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix} + x_2 \begin{bmatrix} 5 \\ -4 \\ -7 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -4 \\ 3 \\ h \end{bmatrix}$$

This vector equation is equivalent to a system of three linear equations in three unknowns. If you row reduce the augmented matrix for this system, you find that

$$\begin{bmatrix} 1 & 5 & -3 & -4 \\ -1 & -4 & 1 & 3 \\ -2 & -7 & 0 & h \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & -3 & -4 \\ 0 & 1 & -2 & -1 \\ 0 & 3 & -6 & h-8 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & -3 & -4 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & 0 & h-5 \end{bmatrix}$$

The system is consistent if and only if there is no pivot in the fourth column. That is, $h - 5$ must be 0. So $\mathbf{y}$ is in Span $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ if and only if $h = 5$.

Remember: The presence of a free variable in a system does not guarantee that the system is consistent.



The points $\begin{bmatrix} -4 \\ 3 \\ h \end{bmatrix}$ lie on a line that intersects the plane when $h = 5$.

1.4 THE MATRIX EQUATION $A\mathbf{x} = \mathbf{b}$

A fundamental idea in linear algebra is to view a linear combination of vectors as the product of a matrix and a vector. The following definition permits us to rephrase some of the concepts of Section 1.3 in new ways.

DEFINITION

If A is an $m \times n$ matrix, with columns $\mathbf{a}_1, \dots, \mathbf{a}_n$, and if $\mathbf{x}$ is in $\mathbb{R}^n$, then the **product of A and $\mathbf{x}$**, denoted by $A\mathbf{x}$, is **the linear combination of the columns of A using the corresponding entries in $\mathbf{x}$ as weights**; that is,

$$A\mathbf{x} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n$$

Note that $A\mathbf{x}$ is defined only if the number of columns of A equals the number of entries in $\mathbf{x}$.

EXAMPLE 1

a. $$\begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 2 \\ -5 \end{bmatrix} + 7\begin{bmatrix} -1 \\ 3 \end{bmatrix}$$
$$= \begin{bmatrix} 4 \\ 0 \end{bmatrix} + \begin{bmatrix} 6 \\ -15 \end{bmatrix} + \begin{bmatrix} -7 \\ 21 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

b. $$\begin{bmatrix} 2 & -3 \\ 8 & 0 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 7 \end{bmatrix} = 4\begin{bmatrix} 2 \\ 8 \\ -5 \end{bmatrix} + 7\begin{bmatrix} -3 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 32 \\ -20 \end{bmatrix} + \begin{bmatrix} -21 \\ 0 \\ 14 \end{bmatrix} = \begin{bmatrix} -13 \\ 32 \\ -6 \end{bmatrix}$$ ■

EXAMPLE 2 For $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ in $\mathbb{R}^m$, write the linear combination $3\mathbf{v}_1 - 5\mathbf{v}_2 + 7\mathbf{v}_3$ as a matrix times a vector.

SOLUTION Place $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ into the columns of a matrix A and place the weights 3, -5, and 7 into a vector $\mathbf{x}$. That is,

$$3\mathbf{v}_1 - 5\mathbf{v}_2 + 7\mathbf{v}_3 = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix} \begin{bmatrix} 3 \\ -5 \\ 7 \end{bmatrix} = A\mathbf{x}$$ ■

Section 1.3 showed how to write a system of linear equations as a vector equation involving a linear combination of vectors. For example, the system

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 4 \\ -5x_2 + 3x_3 &= 1 \end{aligned} \tag{1}$$

is equivalent to

$$x_1\begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2\begin{bmatrix} 2 \\ -5 \end{bmatrix} + x_3\begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix} \tag{2}$$

As in Example 2, the linear combination on the left side is a matrix times a vector, so that (2) becomes

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix} \tag{3}$$

Equation (3) has the form $A\mathbf{x} = \mathbf{b}$. Such an equation is called a **matrix equation**, to distinguish it from a vector equation such as is shown in (2).

Notice how the matrix in (3) is just the matrix of coefficients of the system (1). Similar calculations show that any system of linear equations, or any vector equation such as (2), can be written as an equivalent matrix equation in the form $A\mathbf{x} = \mathbf{b}$. This simple observation will be used repeatedly throughout the text.

Here is the formal result.

THEOREM 3

If A is an $m \times n$ matrix, with columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$, and if $\mathbf{b}$ is in $\mathbb{R}^m$, the matrix equation

$$A\mathbf{x} = \mathbf{b} \tag{4}$$

has the same solution set as the vector equation

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{b} \tag{5}$$

which, in turn, has the same solution set as the system of linear equations whose augmented matrix is

$$\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n & \mathbf{b} \end{bmatrix} \tag{6}$$

Theorem 3 provides a powerful tool for gaining insight into problems in linear algebra, because a system of linear equations may now be viewed in three different but equivalent ways: as a matrix equation, as a vector equation, or as a system of linear equations. Whenever you construct a mathematical model of a problem in real life, you are free to choose whichever viewpoint is most natural. Then you may switch from one formulation of a problem to another whenever it is convenient. In any case, the matrix equation (4), the vector equation (5), and the system of equations are all solved in the same way—by row reducing the augmented matrix (6). Other methods of solution will be discussed later.

Existence of Solutions

The definition of $A\mathbf{x}$ leads directly to the following useful fact.

The equation $A\mathbf{x} = \mathbf{b}$ has a solution if and only if $\mathbf{b}$ is a linear combination of the columns of A.

Section 1.3 considered the existence question, "Is $\mathbf{b}$ in Span $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$?" Equivalently, "Is $A\mathbf{x} = \mathbf{b}$ consistent?" A harder existence problem is to determine whether the equation $A\mathbf{x} = \mathbf{b}$ is consistent *for all* possible $\mathbf{b}$.

EXAMPLE 3 Let $A = \begin{bmatrix} 1 & 3 & 4 \\ -4 & 2 & -6 \\ -3 & -2 & -7 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$. Is the equation $A\mathbf{x} = \mathbf{b}$ consistent for all possible b_1, b_2, b_3?

SOLUTION Row reduce the augmented matrix for $A\mathbf{x} = \mathbf{b}$:

$$\begin{bmatrix} 1 & 3 & 4 & b_1 \\ -4 & 2 & -6 & b_2 \\ -3 & -2 & -7 & b_3 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 4 & b_1 \\ 0 & 14 & 10 & b_2 + 4b_1 \\ 0 & 7 & 5 & b_3 + 3b_1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 3 & 4 & b_1 \\ 0 & 14 & 10 & b_2 + 4b_1 \\ 0 & 0 & 0 & b_3 + 3b_1 - \frac{1}{2}(b_2 + 4b_1) \end{bmatrix}$$

The third entry in column 4 equals $b_1 - \frac{1}{2}b_2 + b_3$. The equation $A\mathbf{x} = \mathbf{b}$ is *not* consistent for every $\mathbf{b}$ because some choices of $\mathbf{b}$ can make $b_1 - \frac{1}{2}b_2 + b_3$ nonzero. ■

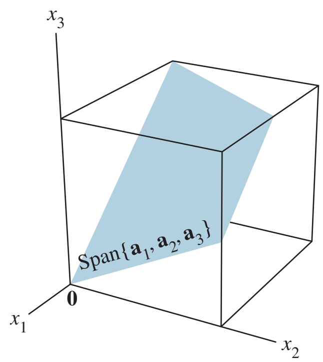


FIGURE 1
The columns of $A = [\,\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3\,]$ span a plane through $\mathbf{0}$.

The reduced matrix in Example 3 provides a description of all $\mathbf{b}$ for which the equation $A\mathbf{x} = \mathbf{b}$ *is* consistent: The entries in $\mathbf{b}$ must satisfy

$$b_1 - \tfrac{1}{2}b_2 + b_3 = 0$$

This is the equation of a plane through the origin in $\mathbb{R}^3$. The plane is the set of all linear combinations of the three columns of A. See Fig. 1.

The equation $A\mathbf{x} = \mathbf{b}$ in Example 3 fails to be consistent for all $\mathbf{b}$ because the echelon form of A has a row of zeros. If A had a pivot in all three rows, we would not care about the calculations in the augmented column because in this case an echelon form of the augmented matrix could not have a row such as $[\,0 \quad 0 \quad 0 \quad 1\,]$.

In the next theorem, the sentence "The columns of A span $\mathbb{R}^m$" means that *every* $\mathbf{b}$ in $\mathbb{R}^m$ is a linear combination of the columns of A. In general, a set of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ in $\mathbb{R}^m$ **spans** (or **generates**) $\mathbb{R}^m$ if every vector in $\mathbb{R}^m$ is a linear combination of $\mathbf{v}_1, \dots, \mathbf{v}_p$—that is, if Span$\{\mathbf{v}_1, \dots, \mathbf{v}_p\} = \mathbb{R}^m$.

THEOREM 4

Let A be an $m \times n$ matrix. Then the following statements are logically equivalent. That is, for a particular A, either they are all true statements or they are all false.

a. For each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.
b. Each $\mathbf{b}$ in $\mathbb{R}^m$ is a linear combination of the columns of A.
c. The columns of A span $\mathbb{R}^m$.
d. A has a pivot position in every row.

Theorem 4 is one of the most useful theorems in this chapter. Statements (a), (b), and (c) are equivalent because of the definition of $A\mathbf{x}$ and what it means for a set of vectors to span $\mathbb{R}^m$. The discussion after Example 3 suggests why (a) and (d) are equivalent; a proof is given at the end of the section. The exercises will provide examples of how Theorem 4 is used.

Warning: Theorem 4 is about a *coefficient matrix*, not an augmented matrix. If an augmented matrix $[\,A \quad \mathbf{b}\,]$ has a pivot position in every row, then the equation $A\mathbf{x} = \mathbf{b}$ may or may not be consistent.

Computation of $A\mathbf{x}$

The calculations in Example 1 were based on the definition of the product of a matrix A and a vector $\mathbf{x}$. The following simple example will lead to a more efficient method for calculating the entries in $A\mathbf{x}$ when working problems by hand.

EXAMPLE 4 Compute $A\mathbf{x}$, where $A = \begin{bmatrix} 2 & 3 & 4 \\ -1 & 5 & -3 \\ 6 & -2 & 8 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$.

SOLUTION From the definition,

$$\begin{bmatrix} 2 & 3 & 4 \\ -1 & 5 & -3 \\ 6 & -2 & 8 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1\begin{bmatrix} 2 \\ -1 \\ 6 \end{bmatrix} + x_2\begin{bmatrix} 3 \\ 5 \\ -2 \end{bmatrix} + x_3\begin{bmatrix} 4 \\ -3 \\ 8 \end{bmatrix}$$

$$= \begin{bmatrix} 2x_1 \\ -x_1 \\ 6x_1 \end{bmatrix} + \begin{bmatrix} 3x_2 \\ 5x_2 \\ -2x_2 \end{bmatrix} + \begin{bmatrix} 4x_3 \\ -3x_3 \\ 8x_3 \end{bmatrix} \tag{7}$$

$$= \begin{bmatrix} 2x_1 + 3x_2 + 4x_3 \\ -x_1 + 5x_2 - 3x_3 \\ 6x_1 - 2x_2 + 8x_3 \end{bmatrix}$$

The first entry in the product $A\mathbf{x}$ is a sum of products (sometimes called a *dot product*), using the first row of A and the entries in $\mathbf{x}$. That is,

$$\begin{bmatrix} 2 & 3 & 4 \\ & & \\ & & \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_1 + 3x_2 + 4x_3 \\ \\ \\ \end{bmatrix}$$

This matrix shows how to compute the first entry in $A\mathbf{x}$ directly, without writing down all the calculations shown in (7). Similarly, the second entry in $A\mathbf{x}$ can be calculated at once by multiplying the entries in the second row of A by the corresponding entries in $\mathbf{x}$ and then summing the resulting products:

$$\begin{bmatrix} & & \\ -1 & 5 & -3 \\ & & \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \\ -x_1 + 5x_2 - 3x_3 \\ \\ \end{bmatrix}$$

Likewise, the third entry in $A\mathbf{x}$ can be calculated from the third row of A and the entries in $\mathbf{x}$. ■

Row–Vector Rule for Computing $A\mathbf{x}$

If the product $A\mathbf{x}$ is defined, then the ith entry in $A\mathbf{x}$ is the sum of the products of corresponding entries from row i of A and from the vector $\mathbf{x}$.

EXAMPLE 5

a. $\begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix}\begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix} = \begin{bmatrix} 1\cdot 4 + 2\cdot 3 + (-1)\cdot 7 \\ 0\cdot 4 + (-5)\cdot 3 + 3\cdot 7 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$

b. $\begin{bmatrix} 2 & -3 \\ 8 & 0 \\ -5 & 2 \end{bmatrix}\begin{bmatrix} 4 \\ 7 \end{bmatrix} = \begin{bmatrix} 2\cdot 4 + (-3)\cdot 7 \\ 8\cdot 4 + 0\cdot 7 \\ (-5)\cdot 4 + 2\cdot 7 \end{bmatrix} = \begin{bmatrix} -13 \\ 32 \\ -6 \end{bmatrix}$

c. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} r \\ s \\ t \end{bmatrix} = \begin{bmatrix} 1\cdot r + 0\cdot s + 0\cdot t \\ 0\cdot r + 1\cdot s + 0\cdot t \\ 0\cdot r + 0\cdot s + 1\cdot t \end{bmatrix} = \begin{bmatrix} r \\ s \\ t \end{bmatrix}$ ■

By definition, the matrix in Example 5(c) with 1's on the diagonal and 0's elsewhere is called an **identity matrix** and is denoted by I. The calculation in part (c) shows that $I\mathbf{x} = \mathbf{x}$ for every $\mathbf{x}$ in $\mathbb{R}^3$. There is an analogous $n \times n$ identity matrix, sometimes written as I_n. As in part (c), $I_n\mathbf{x} = \mathbf{x}$ for every $\mathbf{x}$ in $\mathbb{R}^n$.

Properties of the Matrix–Vector Product $A\mathbf{x}$

The facts in the next theorem are important and will be used throughout the text. The proof relies on the definition of $A\mathbf{x}$ and the algebraic properties of $\mathbb{R}^n$.

THEOREM 5

If A is an $m \times n$ matrix, $\mathbf{u}$ and $\mathbf{v}$ are vectors in $\mathbb{R}^n$, and c is a scalar, then:

a. $A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v}$;

b. $A(c\mathbf{u}) = c(A\mathbf{u})$.

PROOF For simplicity, take $n = 3$, $A = [\,\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3\,]$, and $\mathbf{u}$, $\mathbf{v}$ in $\mathbb{R}^3$. (The proof of the general case is similar.) For $i = 1, 2, 3$, let u_i and v_i be the ith entries in $\mathbf{u}$ and $\mathbf{v}$, respectively. To prove statement (a), compute $A(\mathbf{u} + \mathbf{v})$ as a linear combination of the columns of A using the entries in $\mathbf{u} + \mathbf{v}$ as weights.

$$A(\mathbf{u}+\mathbf{v}) = [\,\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3\,]\begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ u_3 + v_3 \end{bmatrix}$$

Entries in $\mathbf{u} + \mathbf{v}$

$$= (u_1 + v_1)\mathbf{a}_1 + (u_2 + v_2)\mathbf{a}_2 + (u_3 + v_3)\mathbf{a}_3$$

Columns of A

$$= (u_1\mathbf{a}_1 + u_2\mathbf{a}_2 + u_3\mathbf{a}_3) + (v_1\mathbf{a}_1 + v_2\mathbf{a}_2 + v_3\mathbf{a}_3)$$

$$= A\mathbf{u} + A\mathbf{v}$$

To prove statement (b), compute $A(c\mathbf{u})$ as a linear combination of the columns of A using the entries in $c\mathbf{u}$ as weights.

$$A(c\mathbf{u}) = [\,\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3\,]\begin{bmatrix} cu_1 \\ cu_2 \\ cu_3 \end{bmatrix} = (cu_1)\mathbf{a}_1 + (cu_2)\mathbf{a}_2 + (cu_3)\mathbf{a}_3$$

$$= c(u_1\mathbf{a}_1) + c(u_2\mathbf{a}_2) + c(u_3\mathbf{a}_3)$$

$$= c(u_1\mathbf{a}_1 + u_2\mathbf{a}_2 + u_3\mathbf{a}_3)$$

$$= c(A\mathbf{u})$$

■

NUMERICAL NOTE

To optimize a computer algorithm to compute $A\mathbf{x}$, the sequence of calculations should involve data stored in contiguous memory locations. The most widely used professional algorithms for matrix computations are written in Fortran, a language that stores a matrix as a set of columns. Such algorithms compute $A\mathbf{x}$ as a linear combination of the columns of A. In contrast, if a program is written in the popular language C, which stores matrices by rows, $A\mathbf{x}$ should be computed via the alternative rule that uses the rows of A.

PROOF OF THEOREM 4 As was pointed out after Theorem 4, statements (a), (b), and (c) are logically equivalent. So, it suffices to show (for an arbitrary matrix A) that (a) and (d) are either both true or both false. That will tie all four statements together.

Let U be an echelon form of A. Given $\mathbf{b}$ in $\mathbb{R}^m$, we can row reduce the augmented matrix $[\,A \quad \mathbf{b}\,]$ to an augmented matrix $[\,U \quad \mathbf{d}\,]$ for some $\mathbf{d}$ in $\mathbb{R}^m$:

$$[\,A \quad \mathbf{b}\,] \sim \cdots \sim [\,U \quad \mathbf{d}\,]$$

If statement (d) is true, then each row of U contains a pivot position and there can be no pivot in the augmented column. So $A\mathbf{x} = \mathbf{b}$ has a solution for any $\mathbf{b}$, and (a) is true. If (d) is false, the last row of U is all zeros. Let $\mathbf{d}$ be any vector with a 1 in its last entry. Then $[\,U \quad \mathbf{d}\,]$ represents an *inconsistent* system. Since row operations are reversible, $[\,U \quad \mathbf{d}\,]$ can be transformed into the form $[\,A \quad \mathbf{b}\,]$. The new system $A\mathbf{x} = \mathbf{b}$ is also inconsistent, and (a) is false. ■

PRACTICE PROBLEMS

1. Let $A = \begin{bmatrix} 1 & 5 & -2 & 0 \\ -3 & 1 & 9 & -5 \\ 4 & -8 & -1 & 7 \end{bmatrix}$, $\mathbf{p} = \begin{bmatrix} 3 \\ -2 \\ 0 \\ -4 \end{bmatrix}$, and $\mathbf{b} = \begin{bmatrix} -7 \\ 9 \\ 0 \end{bmatrix}$. It can be shown that $\mathbf{p}$ is a solution of $A\mathbf{x} = \mathbf{b}$. Use this fact to exhibit $\mathbf{b}$ as a specific linear combination of the columns of A.

2. Let $A = \begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$, and $\mathbf{v} = \begin{bmatrix} -3 \\ 5 \end{bmatrix}$. Verify Theorem 5(a) in this case by computing $A(\mathbf{u} + \mathbf{v})$ and $A\mathbf{u} + A\mathbf{v}$.

1.4 EXERCISES

Compute the products in Exercises 1–4 using (a) the definition, as in Example 1, and (b) the row–vector rule for computing $A\mathbf{x}$. If a product is undefined, explain why.

1. $\begin{bmatrix} -4 & 2 \\ 1 & 6 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \\ 7 \end{bmatrix}$ **2.** $\begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} \begin{bmatrix} 5 \\ -1 \end{bmatrix}$

3. $\begin{bmatrix} 1 & 2 \\ -3 & 1 \\ 1 & 6 \end{bmatrix} \begin{bmatrix} -2 \\ 3 \end{bmatrix}$ **4.** $\begin{bmatrix} 1 & 3 & -4 \\ 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$

In Exercises 5–8, use the definition of $A\mathbf{x}$ to write the matrix equation as a vector equation, or vice versa.

5. $\begin{bmatrix} 1 & 2 & -3 & 1 \\ -2 & -3 & 1 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -4 \\ 1 \end{bmatrix}$

6. $\begin{bmatrix} 2 & -3 \\ 3 & 2 \\ 8 & -5 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} -3 \\ 5 \end{bmatrix} = \begin{bmatrix} -21 \\ 1 \\ -49 \\ 11 \end{bmatrix}$

7. $x_1 \begin{bmatrix} 4 \\ -1 \\ 7 \\ -4 \end{bmatrix} + x_2 \begin{bmatrix} -5 \\ 3 \\ -5 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} 7 \\ -8 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ -8 \\ 0 \\ -7 \end{bmatrix}$

8. $z_1 \begin{bmatrix} 2 \\ -4 \end{bmatrix} + z_2 \begin{bmatrix} -1 \\ 5 \end{bmatrix} + z_3 \begin{bmatrix} -4 \\ 3 \end{bmatrix} + z_4 \begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 12 \end{bmatrix}$

In Exercises 9 and 10, write the system first as a vector equation and then as a matrix equation.

9. $\begin{aligned} 5x_1 + x_2 - 3x_3 &= 8 \\ 2x_2 + 4x_3 &= 0 \end{aligned}$

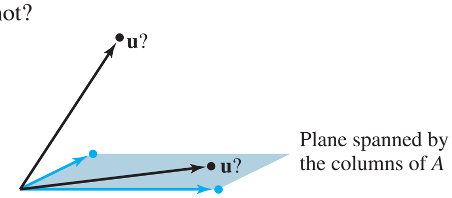
10. $\begin{aligned} 4x_1 - x_2 &= 8 \\ 5x_1 + 3x_2 &= 2 \\ 3x_1 - x_2 &= 1 \end{aligned}$

Given A and $\mathbf{b}$ in Exercises 11 and 12, write the augmented matrix for the linear system that corresponds to the matrix equation $A\mathbf{x} = \mathbf{b}$. Then solve the system and write the solution as a vector.

11. $A = \begin{bmatrix} 1 & 3 & -4 \\ 1 & 5 & 2 \\ -3 & -7 & 6 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} -2 \\ 4 \\ 12 \end{bmatrix}$

12. $A = \begin{bmatrix} 1 & 2 & -1 \\ -3 & -4 & 2 \\ 5 & 2 & 3 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}$

13. Let $\mathbf{u} = \begin{bmatrix} 0 \\ 4 \\ 4 \end{bmatrix}$ and $A = \begin{bmatrix} 3 & -5 \\ -2 & 6 \\ 1 & 1 \end{bmatrix}$. Is $\mathbf{u}$ in the plane in $\mathbb{R}^3$ spanned by the columns of A? (See the figure.) Why or why not?



Where is $\mathbf{u}$?

14. Let $\mathbf{u} = \begin{bmatrix} 4 \\ -1 \\ 4 \end{bmatrix}$ and $A = \begin{bmatrix} 2 & 5 & -1 \\ 0 & 1 & -1 \\ 1 & 2 & 0 \end{bmatrix}$. Is $\mathbf{u}$ in the subset of $\mathbb{R}^3$ spanned by the columns of A? Why or why not?

15. Let $A = \begin{bmatrix} 3 & -1 \\ -9 & 3 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$. Show that the equation $A\mathbf{x} = \mathbf{b}$ does not have a solution for all possible $\mathbf{b}$, and describe the set of all $\mathbf{b}$ for which $A\mathbf{x} = \mathbf{b}$ *does* have a solution.

16. Repeat the requests from Exercise 15 with

$$A = \begin{bmatrix} 1 & -2 & -1 \\ -2 & 2 & 0 \\ 4 & -1 & 3 \end{bmatrix}, \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}.$$

Exercises 17–20 refer to the matrices A and B below. Make appropriate calculations that justify your answers and mention an appropriate theorem.

$$A = \begin{bmatrix} 1 & 3 & 0 & 3 \\ -1 & -1 & -1 & 1 \\ 0 & -4 & 2 & -8 \\ 2 & 0 & 3 & -1 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 4 & 1 & 2 \\ 0 & 1 & 3 & -4 \\ 0 & 2 & 6 & 7 \\ 2 & 9 & 5 & -7 \end{bmatrix}$$

17. How many rows of A contain a pivot position? Does the equation $A\mathbf{x} = \mathbf{b}$ have a solution for each $\mathbf{b}$ in $\mathbb{R}^4$?

18. Can every vector in $\mathbb{R}^4$ be written as a linear combination of the columns of the matrix B above? Do the columns of B span $\mathbb{R}^3$?

19. Can each vector in $\mathbb{R}^4$ be written as a linear combination of the columns of the matrix A above? Do the columns of A span $\mathbb{R}^4$?

20. Do the columns of B span $\mathbb{R}^4$? Does the equation $B\mathbf{x} = \mathbf{y}$ have a solution for each $\mathbf{y}$ in $\mathbb{R}^4$?

21. Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$. Does $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ span $\mathbb{R}^4$? Why or why not?

22. Let $\mathbf{v}_1 = \begin{bmatrix} 0 \\ 0 \\ -3 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 0 \\ -3 \\ 9 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 4 \\ -2 \\ -6 \end{bmatrix}$. Does $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ span $\mathbb{R}^3$? Why or why not?

In Exercises 23 and 24, mark each statement True or False. Justify each answer.

23. a. The equation $A\mathbf{x} = \mathbf{b}$ is referred to as a *vector equation*.

b. A vector $\mathbf{b}$ is a linear combination of the columns of a matrix A if and only if the equation $A\mathbf{x} = \mathbf{b}$ has at least one solution.

c. The equation $A\mathbf{x} = \mathbf{b}$ is consistent if the augmented matrix $[\, A \quad \mathbf{b} \,]$ has a pivot position in every row.

d. The first entry in the product $A\mathbf{x}$ is a sum of products.

e. If the columns of an $m \times n$ matrix A span $\mathbb{R}^m$, then the equation $A\mathbf{x} = \mathbf{b}$ is consistent for each $\mathbf{b}$ in $\mathbb{R}^m$.

f. If A is an $m \times n$ matrix and if the equation $A\mathbf{x} = \mathbf{b}$ is inconsistent for some $\mathbf{b}$ in $\mathbb{R}^m$, then A cannot have a pivot position in every row.

24. a. Every matrix equation $A\mathbf{x} = \mathbf{b}$ corresponds to a vector equation with the same solution set.

b. If the equation $A\mathbf{x} = \mathbf{b}$ is consistent, then $\mathbf{b}$ is in the set spanned by the columns of A.

c. Any linear combination of vectors can always be written in the form $A\mathbf{x}$ for a suitable matrix A and vector $\mathbf{x}$.

d. If the coefficient matrix A has a pivot position in every row, then the equation $A\mathbf{x} = \mathbf{b}$ is inconsistent.

e. The solution set of a linear system whose augmented matrix is $[\, \mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3 \quad \mathbf{b} \,]$ is the same as the solution set of $A\mathbf{x} = \mathbf{b}$, if $A = [\, \mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3 \,]$.

f. If A is an $m \times n$ matrix whose columns do not span $\mathbb{R}^m$, then the equation $A\mathbf{x} = \mathbf{b}$ is consistent for every $\mathbf{b}$ in $\mathbb{R}^m$.

25. Note that $\begin{bmatrix} 4 & -3 & 1 \\ 5 & -2 & 5 \\ -6 & 2 & -3 \end{bmatrix} \begin{bmatrix} -3 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} -7 \\ -3 \\ 10 \end{bmatrix}$. Use this fact (and no row operations) to find scalars c_1, c_2, c_3 such that $\begin{bmatrix} -7 \\ -3 \\ 10 \end{bmatrix} = c_1 \begin{bmatrix} 4 \\ 5 \\ -6 \end{bmatrix} + c_2 \begin{bmatrix} -3 \\ -2 \\ 2 \end{bmatrix} + c_3 \begin{bmatrix} 1 \\ 5 \\ -3 \end{bmatrix}$.

26. Let $\mathbf{u} = \begin{bmatrix} 7 \\ 2 \\ 5 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 3 \\ 1 \\ 3 \end{bmatrix}$, and $\mathbf{w} = \begin{bmatrix} 5 \\ 1 \\ 1 \end{bmatrix}$. It can be shown that $2\mathbf{u} - 3\mathbf{v} - \mathbf{w} = \mathbf{0}$. Use this fact (and no row operations) to find x_1 and x_2 that satisfy the equation $\begin{bmatrix} 7 & 3 \\ 2 & 1 \\ 5 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ 1 \end{bmatrix}$.

27. Rewrite the (numerical) matrix equation below in symbolic form as a vector equation, using symbols $\mathbf{v}_1, \mathbf{v}_2, \ldots$ for the vectors and $c_1, c_2, \ldots$ for scalars. Define what each symbol represents, using the data given in the matrix equation.

$$\begin{bmatrix} -3 & 5 & -4 & 9 & 7 \\ 5 & 8 & 1 & -2 & -4 \end{bmatrix} \begin{bmatrix} -3 \\ 1 \\ 2 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 11 \\ -11 \end{bmatrix}$$

28. Let $\mathbf{q}_1$, $\mathbf{q}_2$, $\mathbf{q}_3$, and $\mathbf{v}$ represent vectors in $\mathbb{R}^5$, and let x_1, x_2, and x_3 denote scalars. Write the following vector equation as a matrix equation. Identify any symbols you choose to use.

$x_1\mathbf{q}_1 + x_2\mathbf{q}_2 + x_3\mathbf{q}_3 = \mathbf{v}$

29. Construct a 3×3 matrix, not in echelon form, whose columns span $\mathbb{R}^3$. Show that the matrix you construct has the desired property.

30. Construct a 3×3 matrix, not in echelon form, whose columns do *not* span $\mathbb{R}^3$. Show that the matrix you construct has the desired property.

31. Let A be a 3×2 matrix. Explain why the equation $A\mathbf{x} = \mathbf{b}$ cannot be consistent for all $\mathbf{b}$ in $\mathbb{R}^3$. Generalize your argument to the case of an arbitrary A with more rows than columns.

32. Could a set of three vectors in $\mathbb{R}^4$ span all of $\mathbb{R}^4$? Explain. What about n vectors in $\mathbb{R}^m$ when n is less than m?

33. Suppose A is a 4×3 matrix and $\mathbf{b}$ is a vector in $\mathbb{R}^4$ with the property that $A\mathbf{x} = \mathbf{b}$ has a unique solution. What can you say about the reduced echelon form of A? Justify your answer.

34. Let A be a 3×4 matrix, let $\mathbf{v}_1$ and $\mathbf{v}_2$ be vectors in $\mathbb{R}^3$, and let $\mathbf{w} = \mathbf{v}_1 + \mathbf{v}_2$. Suppose $\mathbf{v}_1 = A\mathbf{u}_1$ and $\mathbf{v}_2 = A\mathbf{u}_2$ for some vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ in $\mathbb{R}^4$. What fact allows you to conclude that the system $A\mathbf{x} = \mathbf{w}$ is consistent? (*Note:* $\mathbf{u}_1$ and $\mathbf{u}_2$ denote vectors, not scalar entries in vectors.)

35. Let A be a 5×3 matrix, let $\mathbf{y}$ be a vector in $\mathbb{R}^3$, and let $\mathbf{z}$ be a vector in $\mathbb{R}^5$. Suppose $A\mathbf{y} = \mathbf{z}$. What fact allows you to conclude that the system $A\mathbf{x} = 5\mathbf{z}$ is consistent?

36. Suppose A is a 4×4 matrix and $\mathbf{b}$ is a vector in $\mathbb{R}^4$ with the property that $A\mathbf{x} = \mathbf{b}$ has a unique solution. Explain why the columns of A must span $\mathbb{R}^4$.

[**M**] In Exercises 37–40, determine if the columns of the matrix span $\mathbb{R}^4$.

37. $\begin{bmatrix} 7 & 2 & -5 & 8 \\ -5 & -3 & 4 & -9 \\ 6 & 10 & -2 & 7 \\ -7 & 9 & 2 & 15 \end{bmatrix}$

38. $\begin{bmatrix} 4 & -5 & -1 & 8 \\ 3 & -7 & -4 & 2 \\ 5 & -6 & -1 & 4 \\ 9 & 1 & 10 & 7 \end{bmatrix}$

39. $\begin{bmatrix} 10 & -7 & 1 & 4 & 6 \\ -8 & 4 & -6 & -10 & -3 \\ -7 & 11 & -5 & -1 & -8 \\ 3 & -1 & 10 & 12 & 12 \end{bmatrix}$

40. $\begin{bmatrix} 5 & 11 & -6 & -7 & 12 \\ -7 & -3 & -4 & 6 & -9 \\ 11 & 5 & 6 & -9 & -3 \\ -3 & 4 & -7 & 2 & 7 \end{bmatrix}$

41. [**M**] Find a column of the matrix in Exercise 39 that can be deleted and yet have the remaining matrix columns still span $\mathbb{R}^4$.

42. [**M**] Find a column of the matrix in Exercise 40 that can be deleted and yet have the remaining matrix columns still span $\mathbb{R}^4$. Can you delete more than one column?

SG Mastering Linear Algebra Concepts: Span 1–18

WEB

SOLUTIONS TO PRACTICE PROBLEMS

1. The matrix equation

$$\begin{bmatrix} 1 & 5 & -2 & 0 \\ -3 & 1 & 9 & -5 \\ 4 & -8 & -1 & 7 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \\ 0 \\ -4 \end{bmatrix} = \begin{bmatrix} -7 \\ 9 \\ 0 \end{bmatrix}$$

is equivalent to the vector equation

$$3\begin{bmatrix} 1 \\ -3 \\ 4 \end{bmatrix} - 2\begin{bmatrix} 5 \\ 1 \\ -8 \end{bmatrix} + 0\begin{bmatrix} -2 \\ 9 \\ -1 \end{bmatrix} - 4\begin{bmatrix} 0 \\ -5 \\ 7 \end{bmatrix} = \begin{bmatrix} -7 \\ 9 \\ 0 \end{bmatrix}$$

which expresses $\mathbf{b}$ as a linear combination of the columns of A.

2.
$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} 4 \\ -1 \end{bmatrix} + \begin{bmatrix} -3 \\ 5 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$$

$$A(\mathbf{u} + \mathbf{v}) = \begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 + 20 \\ 3 + 4 \end{bmatrix} = \begin{bmatrix} 22 \\ 7 \end{bmatrix}$$

$$\begin{aligned} A\mathbf{u} + A\mathbf{v} &= \begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 4 \\ -1 \end{bmatrix} + \begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} -3 \\ 5 \end{bmatrix} \\ &= \begin{bmatrix} 3 \\ 11 \end{bmatrix} + \begin{bmatrix} 19 \\ -4 \end{bmatrix} = \begin{bmatrix} 22 \\ 7 \end{bmatrix} \end{aligned}$$

2 Matrix Algebra

INTRODUCTORY EXAMPLE

Computer Models in Aircraft Design

To design the next generation of commercial and military aircraft, engineers at Boeing's Phantom Works use 3D modeling and computational fluid dynamics (CFD). They study the airflow around a virtual airplane to answer important design questions before physical models are created. This has drastically reduced design cycle times and cost—and linear algebra plays a crucial role in the process.

The virtual airplane begins as a mathematical "wire-frame" model that exists only in computer memory and on graphics display terminals. (A model of a Boeing 777 is shown.) This mathematical model organizes and influences each step of the design and manufacture of the airplane—both the exterior and interior. The CFD analysis concerns the exterior surface.

Although the finished skin of a plane may seem smooth, the geometry of the surface is complicated. In addition to wings and a fuselage, an aircraft has nacelles, stabilizers, slats, flaps, and ailerons. The way air flows around these structures determines how the plane moves through the sky. Equations that describe the airflow are complicated, and they must account for engine intake, engine exhaust, and the wakes left by the wings of the plane. To study the airflow, engineers need a highly refined description of the plane's surface.

A computer creates a model of the surface by first superimposing a three-dimensional grid of "boxes" on the original wire-frame model. Boxes in this grid lie either completely inside or completely outside the plane, or they intersect the surface of the plane. The computer selects the boxes that intersect the surface and subdivides them, retaining only the smaller boxes that still intersect the surface. The subdividing process is repeated until the grid is extremely fine. A typical grid can include over 400,000 boxes.

The process for finding the airflow around the plane involves repeatedly solving a system of linear equations $A\mathbf{x} = \mathbf{b}$ that may involve up to 2 million equations and variables. The vector $\mathbf{b}$ changes each time, based on data from the grid and solutions of previous equations. Using the fastest computers available commercially, a Phantom Works team can spend from a few hours to several days setting up and solving a single airflow problem. After the team analyzes the solution, they may make small changes to the airplane surface and begin the whole process again. Thousands of CFD runs may be required.

This chapter presents two important concepts that assist in the solution of such massive systems of equations:

- *Partitioned matrices:* A typical CFD system of equations has a "sparse" coefficient matrix with mostly zero entries. Grouping the variables correctly leads to a partitioned matrix with many zero blocks. Section 2.4 introduces such matrices and describes some of their applications.

- *Matrix factorizations:* Even when written with partitioned matrices, the system of equations is complicated. To further simplify the computations, the CFD software at Boeing uses what is called an LU factorization of the coefficient matrix. Section 2.5 discusses LU and other useful matrix factorizations. Further details about factorizations appear at several points later in the text.

To analyze a solution of an airflow system, engineers want to visualize the airflow over the surface of the plane. They use computer graphics, and linear algebra provides the engine for the graphics. The wire-frame model of the plane's surface is stored as data in many matrices. Once the image has been rendered on a computer screen, engineers can change its scale, zoom in or out of small regions, and rotate the image to see parts that may be hidden from view. Each of these operations is accomplished by appropriate matrix multiplications. Section 2.7 explains the basic ideas.

Modern CFD has revolutionized wing design. The Boeing Blended Wing Body is in design for the year 2020 or sooner.

WEB

Our ability to analyze and solve equations will be greatly enhanced when we can perform algebraic operations with matrices. Furthermore, the definitions and theorems in this chapter provide some basic tools for handling the many applications of linear algebra that involve two or more matrices. For square matrices, the Invertible Matrix Theorem in Section 2.3 ties together most of the concepts treated earlier in the text. Sections 2.4 and 2.5 examine partitioned matrices and matrix factorizations, which appear in most modern uses of linear algebra. Sections 2.6 and 2.7 describe two interesting applications of matrix algebra, to economics and to computer graphics.

2.1 MATRIX OPERATIONS

If A is an $m \times n$ matrix—that is, a matrix with m rows and n columns—then the scalar entry in the ith row and jth column of A is denoted by a_{ij} and is called the (i, j)-entry of A. See Fig. 1. For instance, the (3, 2)-entry is the number a_{32} in the third row, second column. Each column of A is a list of m real numbers, which identifies a vector in $\mathbb{R}^m$. Often, these columns are denoted by $\mathbf{a}_1, \dots, \mathbf{a}_n$, and the matrix A is written as

$$A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}$$

Observe that the number a_{ij} is the ith entry (from the top) of the jth column vector $\mathbf{a}_j$.

The **diagonal entries** in an $m \times n$ matrix $A = [\,a_{ij}\,]$ are $a_{11}, a_{22}, a_{33}, \dots$, and they form the **main diagonal** of A. A **diagonal matrix** is a square $n \times n$ matrix whose nondiagonal entries are zero. An example is the $n \times n$ identity matrix, I_n. An $m \times n$ matrix whose entries are all zero is a **zero matrix** and is written as 0. The size of a zero matrix is usually clear from the context.

$$
\begin{array}{rc}
 & \begin{array}{ccccc} & & \text{Column} & & \\ & & j & & \end{array} \\
\text{Row } i & \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix} = A \\
 & \begin{array}{ccccc} \uparrow & & \uparrow & & \uparrow \\ \mathbf{a}_1 & & \mathbf{a}_j & & \mathbf{a}_n \end{array}
\end{array}
$$

FIGURE 1 Matrix notation.

Sums and Scalar Multiples

The arithmetic for vectors described earlier has a natural extension to matrices. We say that two matrices are **equal** if they have the same size (i.e., the same number of rows and the same number of columns) and if their corresponding columns are equal, which amounts to saying that their corresponding entries are equal. If A and B are $m \times n$ matrices, then the **sum** $A + B$ is the $m \times n$ matrix whose columns are the sums of the corresponding columns in A and B. Since vector addition of the columns is done entrywise, each entry in $A + B$ is the sum of the corresponding entries in A and B. The sum $A + B$ is defined only when A and B are the same size.

EXAMPLE 1 Let

$$A = \begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 7 \end{bmatrix}, \quad C = \begin{bmatrix} 2 & -3 \\ 0 & 1 \end{bmatrix}$$

Then

$$A + B = \begin{bmatrix} 5 & 1 & 6 \\ 2 & 8 & 9 \end{bmatrix}$$

but $A + C$ is not defined because A and C have different sizes. ∎

If r is a scalar and A is a matrix, then the **scalar multiple** rA is the matrix whose columns are r times the corresponding columns in A. As with vectors, $-A$ stands for $(-1)A$, and $A - B$ is the same as $A + (-1)B$.

EXAMPLE 2 If A and B are the matrices in Example 1, then

$$2B = 2\begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 7 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 2 \\ 6 & 10 & 14 \end{bmatrix}$$

$$A - 2B = \begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix} - \begin{bmatrix} 2 & 2 & 2 \\ 6 & 10 & 14 \end{bmatrix} = \begin{bmatrix} 2 & -2 & 3 \\ -7 & -7 & -12 \end{bmatrix}$$ ∎

It was unnecessary in Example 2 to compute $A - 2B$ as $A + (-1)2B$ because the usual rules of algebra apply to sums and scalar multiples of matrices, as the following theorem shows.

THEOREM 1

Let A, B, and C be matrices of the same size, and let r and s be scalars.

a. $A + B = B + A$
b. $(A + B) + C = A + (B + C)$
c. $A + 0 = A$
d. $r(A + B) = rA + rB$
e. $(r + s)A = rA + sA$
f. $r(sA) = (rs)A$

Each equality in Theorem 1 is verified by showing that the matrix on the left side has the same size as the matrix on the right and that corresponding columns are equal. Size is no problem because A, B, and C are equal in size. The equality of columns follows immediately from analogous properties of vectors. For instance, if the jth columns of A, B, and C are $\mathbf{a}_j$, $\mathbf{b}_j$, and $\mathbf{c}_j$, respectively, then the jth columns of $(A+B)+C$ and $A+(B+C)$ are

$$(\mathbf{a}_j+\mathbf{b}_j)+\mathbf{c}_j \quad \text{and} \quad \mathbf{a}_j+(\mathbf{b}_j+\mathbf{c}_j)$$

respectively. Since these two vector sums are equal for each j, property (b) is verified.

Because of the associative property of addition, we can simply write $A+B+C$ for the sum, which can be computed either as $(A+B)+C$ or as $A+(B+C)$. The same applies to sums of four or more matrices.

Matrix Multiplication

When a matrix B multiplies a vector $\mathbf{x}$, it transforms $\mathbf{x}$ into the vector $B\mathbf{x}$. If this vector is then multiplied in turn by a matrix A, the resulting vector is $A(B\mathbf{x})$. See Fig. 2.

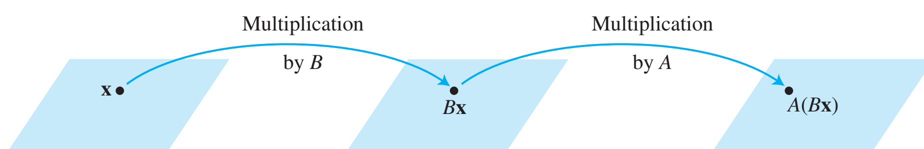


FIGURE 2 Multiplication by B and then A.

Thus $A(B\mathbf{x})$ is produced from $\mathbf{x}$ by a *composition* of mappings—the linear transformations studied in Section 1.8. Our goal is to represent this composite mapping as multiplication by a single matrix, denoted by AB, so that

$$A(B\mathbf{x}) = (AB)\mathbf{x} \tag{1}$$

See Fig. 3.

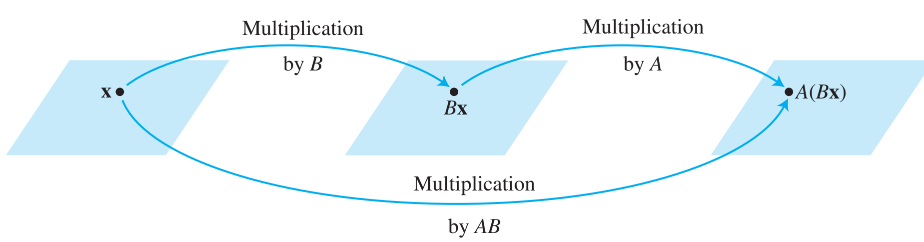


FIGURE 3 Multiplication by AB.

If A is $m\times n$, B is $n\times p$, and $\mathbf{x}$ is in $\mathbb{R}^p$, denote the columns of B by $\mathbf{b}_1,\dots,\mathbf{b}_p$ and the entries in $\mathbf{x}$ by $x_1,\dots,x_p$. Then

$$B\mathbf{x} = x_1\mathbf{b}_1+\cdots+x_p\mathbf{b}_p$$

By the linearity of multiplication by A,

$$\begin{aligned} A(B\mathbf{x}) &= A(x_1\mathbf{b}_1)+\cdots+A(x_p\mathbf{b}_p) \\ &= x_1A\mathbf{b}_1+\cdots+x_pA\mathbf{b}_p \end{aligned}$$

The vector $A(B\mathbf{x})$ is a linear combination of the vectors $A\mathbf{b}_1, \dots, A\mathbf{b}_p$, using the entries in $\mathbf{x}$ as weights. In matrix notation, this linear combination is written as

$$A(B\mathbf{x}) = [\, A\mathbf{b}_1 \quad A\mathbf{b}_2 \quad \cdots \quad A\mathbf{b}_p \,]\mathbf{x}$$

Thus multiplication by $[\, A\mathbf{b}_1 \quad A\mathbf{b}_2 \quad \cdots \quad A\mathbf{b}_p \,]$ transforms $\mathbf{x}$ into $A(B\mathbf{x})$. We have found the matrix we sought!

DEFINITION

If A is an $m \times n$ matrix, and if B is an $n \times p$ matrix with columns $\mathbf{b}_1, \dots, \mathbf{b}_p$, then the product AB is the $m \times p$ matrix whose columns are $A\mathbf{b}_1, \dots, A\mathbf{b}_p$. That is,

$$AB = A\begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_p \end{bmatrix} = \begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_p \end{bmatrix}$$

This definition makes equation (1) true for all $\mathbf{x}$ in $\mathbb{R}^p$. Equation (1) proves that the composite mapping in Fig. 3 is a linear transformation and that its standard matrix is AB. *Multiplication of matrices corresponds to composition of linear transformations.*

EXAMPLE 3 Compute AB, where $A = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix}$.

SOLUTION Write $B = [\, \mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3 \,]$, and compute:

$$A\mathbf{b}_1 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \end{bmatrix}, \quad A\mathbf{b}_2 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 3 \\ -2 \end{bmatrix}, \quad A\mathbf{b}_3 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 6 \\ 3 \end{bmatrix}$$

$$= \begin{bmatrix} 11 \\ -1 \end{bmatrix} \qquad\qquad = \begin{bmatrix} 0 \\ 13 \end{bmatrix} \qquad\qquad = \begin{bmatrix} 21 \\ -9 \end{bmatrix}$$

Then

$$AB = A[\, \mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3 \,] = \begin{array}{c} \begin{bmatrix} 11 & 0 & 21 \\ -1 & 13 & -9 \end{bmatrix} \\ \begin{array}{ccc} \uparrow & \uparrow & \uparrow \\ A\mathbf{b}_1 & A\mathbf{b}_2 & A\mathbf{b}_3 \end{array} \end{array}$$

■

Notice that since the first column of AB is $A\mathbf{b}_1$, this column is a linear combination of the columns of A using the entries in $\mathbf{b}_1$ as weights. A similar statement is true for each column of AB.

Each column of AB is a linear combination of the columns of A using weights from the corresponding column of B.

Obviously, the number of columns of A must match the number of rows in B in order for a linear combination such as $A\mathbf{b}_1$ to be defined. Also, the definition of AB shows that *AB has the same number of rows as A and the same number of columns as B.*

EXAMPLE 4 If A is a 3×5 matrix and B is a 5×2 matrix, what are the sizes of AB and BA, if they are defined?

SOLUTION Since A has 5 columns and B has 5 rows, the product AB is defined and is a 3×2 matrix:

$$
\underset{3\times 5}{\overset{A}{\begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix}}}
\underset{5\times 2}{\overset{B}{\begin{bmatrix} * & * \\ * & * \\ * & * \\ * & * \\ * & * \end{bmatrix}}}
=
\underset{3\times 2}{\overset{AB}{\begin{bmatrix} * & * \\ * & * \\ * & * \end{bmatrix}}}
$$

Match

Size of AB

The product BA is *not* defined because the 2 columns of B do not match the 3 rows of A. ■

The definition of AB is important for theoretical work and applications, but the following rule provides a more efficient method for calculating the individual entries in AB when working small problems by hand.

ROW–COLUMN RULE FOR COMPUTING AB

If the product AB is defined, then the entry in row i and column j of AB is the sum of the products of corresponding entries from row i of A and column j of B. If $(AB)_{ij}$ denotes the (i, j)-entry in AB, and if A is an $m \times n$ matrix, then

$$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

To verify this rule, let $B = [\,\mathbf{b}_1 \ \ \cdots \ \ \mathbf{b}_p\,]$. Column j of AB is $A\mathbf{b}_j$, and we can compute $A\mathbf{b}_j$ by the row–vector rule for computing $A\mathbf{x}$ from Section 1.4. The ith entry in $A\mathbf{b}_j$ is the sum of the products of corresponding entries from row i of A and the vector $\mathbf{b}_j$, which is precisely the computation described in the rule for computing the (i, j)-entry of AB.

EXAMPLE 5 Use the row–column rule to compute two of the entries in AB for the matrices in Example 3. An inspection of the numbers involved will make it clear how the two methods for calculating AB produce the same matrix.

SOLUTION To find the entry in row 1 and column 3 of AB, consider row 1 of A and column 3 of B. Multiply corresponding entries and add the results, as shown below:

$$AB = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix} = \begin{bmatrix} \square & \square & 2(6) + 3(3) \\ \square & \square & \square \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & \square & \square \end{bmatrix}$$

For the entry in row 2 and column 2 of AB, use row 2 of A and column 2 of B:

$$\begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & 1(3) + -5(-2) & \square \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & 13 & \square \end{bmatrix}$$ ■

EXAMPLE 6 Find the entries in the second row of AB, where

$$A = \begin{bmatrix} 2 & -5 & 0 \\ -1 & 3 & -4 \\ 6 & -8 & -7 \\ -3 & 0 & 9 \end{bmatrix}, \qquad B = \begin{bmatrix} 4 & -6 \\ 7 & 1 \\ 3 & 2 \end{bmatrix}$$

SOLUTION By the row–column rule, the entries of the second row of AB come from row 2 of A (and the columns of B):

$$\begin{bmatrix} 2 & -5 & 0 \\ -1 & 3 & -4 \\ 6 & -8 & -7 \\ -3 & 0 & 9 \end{bmatrix} \begin{bmatrix} 4 & -6 \\ 7 & 1 \\ 3 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} \square & \square \\ -4 + 21 - 12 & 6 + 3 - 8 \\ \square & \square \\ \square & \square \end{bmatrix} = \begin{bmatrix} \square & \square \\ 5 & 1 \\ \square & \square \\ \square & \square \end{bmatrix}$$

Notice that since Example 6 requested only the second row of AB, we could have written just the second row of A to the left of B and computed

$$\begin{bmatrix} -1 & 3 & -4 \end{bmatrix} \begin{bmatrix} 4 & -6 \\ 7 & 1 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 1 \end{bmatrix}$$

This observation about rows of AB is true in general and follows from the row–column rule. Let $\text{row}_i(A)$ denote the ith row of a matrix A. Then

$$\text{row}_i(AB) = \text{row}_i(A) \cdot B \tag{2}$$

Properties of Matrix Multiplication

The following theorem lists the standard properties of matrix multiplication. Recall that I_m represents the $m \times m$ identity matrix and $I_m\mathbf{x} = \mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^m$.

THEOREM 2

Let A be an $m \times n$ matrix, and let B and C have sizes for which the indicated sums and products are defined.

a. $A(BC) = (AB)C$ (associative law of multiplication)

b. $A(B + C) = AB + AC$ (left distributive law)

c. $(B + C)A = BA + CA$ (right distributive law)

d. $r(AB) = (rA)B = A(rB)$ for any scalar r

e. $I_mA = A = AI_n$ (identity for matrix multiplication)

PROOF Properties (b)–(e) are considered in the exercises. Property (a) follows from the fact that matrix multiplication corresponds to composition of linear transformations (which are functions), and it is known (or easy to check) that the composition of functions is associative. Here is another proof of (a) that rests on the "column definition" of the product of two matrices. Let

$$C = [\,\mathbf{c}_1 \quad \cdots \quad \mathbf{c}_p\,]$$

By the definition of matrix multiplication,

$$BC = [\, B\mathbf{c}_1 \quad \cdots \quad B\mathbf{c}_p \,]$$
$$A(BC) = [\, A(B\mathbf{c}_1) \quad \cdots \quad A(B\mathbf{c}_p) \,]$$

Recall from equation (1) that the definition of AB makes $A(B\mathbf{x}) = (AB)\mathbf{x}$ for all $\mathbf{x}$, so

$$A(BC) = [\, (AB)\mathbf{c}_1 \quad \cdots \quad (AB)\mathbf{c}_p \,] = (AB)C$$

The associative and distributive laws in Theorems 1 and 2 say essentially that pairs of parentheses in matrix expressions can be inserted and deleted in the same way as in the algebra of real numbers. In particular, we can write ABC for the product, which can be computed either as $A(BC)$ or as $(AB)C$.[1] Similarly, a product $ABCD$ of four matrices can be computed as $A(BCD)$ or $(ABC)D$ or $A(BC)D$, and so on. It does not matter how we group the matrices when computing the product, so long as the left-to-right order of the matrices is preserved.

The left-to-right order in products is critical because AB and BA are usually not the same. This is not surprising, because the columns of AB are linear combinations of the columns of A, whereas the columns of BA are constructed from the columns of B. The position of the factors in the product AB is emphasized by saying that A is *right-multiplied* by B or that B is *left-multiplied* by A. If $AB = BA$, we say that A and B **commute** with one another.

EXAMPLE 7 Let $A = \begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix}$. Show that these matrices do not commute. That is, verify that $AB \neq BA$.

SOLUTION

$$AB = \begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 14 & 3 \\ -2 & -6 \end{bmatrix}$$
$$BA = \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix}\begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix} = \begin{bmatrix} 10 & 2 \\ 29 & -2 \end{bmatrix}$$

Example 7 illustrates the first of the following list of important differences between matrix algebra and the ordinary algebra of real numbers. See Exercises 9–12 for examples of these situations.

WARNINGS:

1. In general, $AB \neq BA$.
2. The cancellation laws do *not* hold for matrix multiplication. That is, if $AB = AC$, then it is *not* true in general that $B = C$. (See Exercise 10.)
3. If a product AB is the zero matrix, you *cannot* conclude in general that either $A = 0$ or $B = 0$. (See Exercise 12.)

Powers of a Matrix

WEB

If A is an $n \times n$ matrix and if k is a positive integer, then A^k denotes the product of k

[1] When B is square and C has fewer columns than A has rows, it is more efficient to compute $A(BC)$ than $(AB)C$.

copies of A:

$$A^k = \underbrace{A \cdots A}_{k}$$

If A is nonzero and if $\mathbf{x}$ is in $\mathbb{R}^n$, then $A^k\mathbf{x}$ is the result of left-multiplying $\mathbf{x}$ by A repeatedly k times. If $k = 0$, then $A^0\mathbf{x}$ should be $\mathbf{x}$ itself. Thus A^0 is interpreted as the identity matrix. Matrix powers are useful in both theory and applications (Sections 2.6, 4.9, and later in the text).

The Transpose of a Matrix

Given an $m \times n$ matrix A, the **transpose** of A is the $n \times m$ matrix, denoted by A^T, whose columns are formed from the corresponding rows of A.

EXAMPLE 8 Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} -5 & 2 \\ 1 & -3 \\ 0 & 4 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -3 & 5 & -2 & 7 \end{bmatrix}$$

Then

$$A^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}, \quad B^T = \begin{bmatrix} -5 & 1 & 0 \\ 2 & -3 & 4 \end{bmatrix}, \quad C^T = \begin{bmatrix} 1 & -3 \\ 1 & 5 \\ 1 & -2 \\ 1 & 7 \end{bmatrix}$$

THEOREM 3

Let A and B denote matrices whose sizes are appropriate for the following sums and products.

a. $(A^T)^T = A$
b. $(A + B)^T = A^T + B^T$
c. For any scalar r, $(rA)^T = rA^T$
d. $(AB)^T = B^TA^T$

Proofs of (a)–(c) are straightforward and are omitted. For (d), see Exercise 33. Usually, $(AB)^T$ is not equal to A^TB^T, even when A and B have sizes such that the product A^TB^T is defined.

The generalization of Theorem 3(d) to products of more than two factors can be stated in words as follows:

The transpose of a product of matrices equals the product of their transposes in the *reverse* order.

The exercises contain numerical examples that illustrate properties of transposes.

NUMERICAL NOTES

1. The fastest way to obtain AB on a computer depends on the way in which the computer stores matrices in its memory. The standard high-performance algorithms, such as in LAPACK, calculate AB by columns, as in our definition of the product. (A version of LAPACK written in C++ calculates AB by rows.)
2. The definition of AB lends itself well to parallel processing on a computer. The columns of B are assigned individually or in groups to different processors, which independently and hence simultaneously compute the corresponding columns of AB.

PRACTICE PROBLEMS

1. Since vectors in $\mathbb{R}^n$ may be regarded as $n \times 1$ matrices, the properties of transposes in Theorem 3 apply to vectors, too. Let

$$A = \begin{bmatrix} 1 & -3 \\ -2 & 4 \end{bmatrix} \quad \text{and} \quad \mathbf{x} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

Compute $(A\mathbf{x})^T$, $\mathbf{x}^T A^T$, $\mathbf{x}\mathbf{x}^T$, and $\mathbf{x}^T\mathbf{x}$. Is $A^T\mathbf{x}^T$ defined?

2. Let A be a 4×4 matrix and let $\mathbf{x}$ be a vector in $\mathbb{R}^4$. What is the fastest way to compute $A^2\mathbf{x}$? Count the multiplications.

2.1 EXERCISES

In Exercises 1 and 2, compute each matrix sum or product if it is defined. If an expression is undefined, explain why. Let

$$A = \begin{bmatrix} 2 & 0 & -1 \\ 4 & -5 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 7 & -5 & 1 \\ 1 & -4 & -3 \end{bmatrix},$$

$$C = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 3 & 5 \\ -1 & 4 \end{bmatrix}, \quad E = \begin{bmatrix} -5 \\ 3 \end{bmatrix}$$

1. $-2A$, $B - 2A$, AC, CD

2. $A + 3B$, $2C - 3E$, DB, EC

In the rest of this exercise set and in those to follow, assume that each matrix expression is defined. That is, the sizes of the matrices (and vectors) involved "match" appropriately.

3. Let $A = \begin{bmatrix} 2 & -5 \\ 3 & -2 \end{bmatrix}$. Compute $3I_2 - A$ and $(3I_2)A$.

4. Compute $A - 5I_3$ and $(5I_3)A$, where

$$A = \begin{bmatrix} 5 & -1 & 3 \\ -4 & 3 & -6 \\ -3 & 1 & 2 \end{bmatrix}.$$

In Exercises 5 and 6, compute the product AB in two ways: (a) by the definition, where $A\mathbf{b}_1$ and $A\mathbf{b}_2$ are computed separately, and (b) by the row–column rule for computing AB.

5. $A = \begin{bmatrix} -1 & 3 \\ 2 & 4 \\ 5 & -3 \end{bmatrix}$, $B = \begin{bmatrix} 4 & -2 \\ -2 & 3 \end{bmatrix}$

6. $A = \begin{bmatrix} 4 & -3 \\ -3 & 5 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 4 \\ 3 & -2 \end{bmatrix}$

7. If a matrix A is 5×3 and the product AB is 5×7, what is the size of B?

8. How many rows does B have if BC is a 5×4 matrix?

9. Let $A = \begin{bmatrix} 2 & 3 \\ -1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 9 \\ -3 & k \end{bmatrix}$. What value(s) of k, if any, will make $AB = BA$?

10. Let $A = \begin{bmatrix} 3 & -6 \\ -1 & 2 \end{bmatrix}$, $B = \begin{bmatrix} -1 & 1 \\ 3 & 4 \end{bmatrix}$, and $C = \begin{bmatrix} -3 & -5 \\ 2 & 1 \end{bmatrix}$. Verify that $AB = AC$ and yet $B \neq C$.

11. Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$ and $D = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}$. Compute AD and DA. Explain how the columns or rows of A change when A is multiplied by D on the right or on the left. Find a 3×3 matrix B, not the identity matrix or the zero matrix, such that $AB = BA$.

12. Let $A = \begin{bmatrix} 3 & -6 \\ -2 & 4 \end{bmatrix}$. Construct a 2×2 matrix B such that AB is the zero matrix. Use two different nonzero columns for B.

13. Let $\mathbf{r}_1, \ldots, \mathbf{r}_p$ be vectors in $\mathbb{R}^n$, and let Q be an $m \times n$ matrix. Write the matrix $[\, Q\mathbf{r}_1 \quad \cdots \quad Q\mathbf{r}_p \,]$ as a *product* of two matrices (neither of which is an identity matrix).

14. Let U be the 3×2 cost matrix described in Example 6 in Section 1.8. The first column of U lists the costs per dollar of output for manufacturing product B, and the second column lists the costs per dollar of output for product C. (The costs are categorized as materials, labor, and overhead.) Let $\mathbf{q}_1$ be a vector in $\mathbb{R}^2$ that lists the output (measured in dollars) of products B and C manufactured during the first quarter of the year, and let $\mathbf{q}_2$, $\mathbf{q}_3$, and $\mathbf{q}_4$ be the analogous vectors that list the amounts of products B and C manufactured in the second, third, and fourth quarters, respectively. Give an economic description of the data in the matrix UQ, where $Q = [\, \mathbf{q}_1 \quad \mathbf{q}_2 \quad \mathbf{q}_3 \quad \mathbf{q}_4 \,]$.

Exercises 15 and 16 concern arbitrary matrices A, B, and C for which the indicated sums and products are defined. Mark each statement True or False. Justify each answer.

15. a. If A and B are 2×2 matrices with columns $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{b}_1$, $\mathbf{b}_2$, respectively, then $AB = [\, \mathbf{a}_1\mathbf{b}_1 \quad \mathbf{a}_2\mathbf{b}_2 \,]$.

b. Each column of AB is a linear combination of the columns of B using weights from the corresponding column of A.

c. $AB + AC = A(B + C)$

d. $A^T + B^T = (A + B)^T$

e. The transpose of a product of matrices equals the product of their transposes in the same order.

16. a. The first row of AB is the first row of A multiplied on the right by B.

b. If A and B are 3×3 matrices and $B = [\, \mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3 \,]$, then $AB = [\, A\mathbf{b}_1 + A\mathbf{b}_2 + A\mathbf{b}_3 \,]$.

c. If A is an $n \times n$ matrix, then $(A^2)^T = (A^T)^2$

d. $(ABC)^T = C^T A^T B^T$

e. The transpose of a sum of matrices equals the sum of their transposes.

17. If $A = \begin{bmatrix} 1 & -3 \\ -3 & 5 \end{bmatrix}$ and $AB = \begin{bmatrix} -3 & -11 \\ 1 & 17 \end{bmatrix}$, determine the first and second columns of B.

18. Suppose the third column of B is all zeros. What can be said about the third column of AB?

19. Suppose the third column of B is the sum of the first two columns. What can be said about the third column of AB? Why?

20. Suppose the first two columns, $\mathbf{b}_1$ and $\mathbf{b}_2$, of B are equal. What can be said about the columns of AB? Why?

21. Suppose the last column of AB is entirely zeros but B itself has no column of zeros. What can be said about the columns of A?

22. Show that if the columns of B are linearly dependent, then so are the columns of AB.

23. Suppose $CA = I_n$ (the $n \times n$ identity matrix). Show that the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution. Explain why A cannot have more columns than rows.

24. Suppose A is a $3 \times n$ matrix whose columns span $\mathbb{R}^3$. Explain how to construct an $n \times 3$ matrix D such that $AD = I_3$.

25. Suppose A is an $m \times n$ matrix and there exist $n \times m$ matrices C and D such that $CA = I_n$ and $AD = I_m$. Prove that $m = n$ and $C = D$. [*Hint:* Think about the product CAD.]

26. Suppose $AD = I_m$ (the $m \times m$ identity matrix). Show that for any $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution. [*Hint:* Think about the equation $AD\mathbf{b} = \mathbf{b}$.] Explain why A cannot have more rows than columns.

In Exercises 27 and 28, view vectors in $\mathbb{R}^n$ as $n \times 1$ matrices. For $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$, the matrix product $\mathbf{u}^T\mathbf{v}$ is a 1×1 matrix, called the **scalar product**, or **inner product**, of $\mathbf{u}$ and $\mathbf{v}$. It is usually written as a single real number without brackets. The matrix product $\mathbf{u}\mathbf{v}^T$ is an $n \times n$ matrix, called the **outer product** of $\mathbf{u}$ and $\mathbf{v}$. The products $\mathbf{u}^T\mathbf{v}$ and $\mathbf{u}\mathbf{v}^T$ will appear later in the text.

27. Let $\mathbf{u} = \begin{bmatrix} -3 \\ 2 \\ -5 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$. Compute $\mathbf{u}^T\mathbf{v}$, $\mathbf{v}^T\mathbf{u}$, $\mathbf{u}\mathbf{v}^T$, and $\mathbf{v}\mathbf{u}^T$.

28. If $\mathbf{u}$ and $\mathbf{v}$ are in $\mathbb{R}^n$, how are $\mathbf{u}^T\mathbf{v}$ and $\mathbf{v}^T\mathbf{u}$ related? How are $\mathbf{u}\mathbf{v}^T$ and $\mathbf{v}\mathbf{u}^T$ related?

29. Prove Theorem 2(b) and 2(c). Use the row–column rule. The (i, j)-entry in $A(B + C)$ can be written as

$$a_{i1}(b_{1j} + c_{1j}) + \cdots + a_{in}(b_{nj} + c_{nj})$$

or

$$\sum_{k=1}^{n} a_{ik}(b_{kj} + c_{kj})$$

30. Prove Theorem 2(d). [*Hint:* The (i, j)-entry in $(rA)B$ is $(ra_{i1})b_{1j} + \cdots + (ra_{in})b_{nj}$.]

31. Show that $I_m A = A$ where A is an $m \times n$ matrix. Assume $I_m\mathbf{x} = \mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^m$.

32. Show that $AI_n = A$ when A is an $m \times n$ matrix. [*Hint:* Use the (column) definition of AI_n.]

33. Prove Theorem 3(d). [*Hint:* Consider the jth row of $(AB)^T$.]

34. Give a formula for $(AB\mathbf{x})^T$, where $\mathbf{x}$ is a vector and A and B are matrices of appropriate sizes.

35. **[M]** Read the documentation for your matrix program, and write the commands that will produce the following matrices (without keying in each entry of the matrix).

a. A 4×5 matrix of zeros

b. A 5×3 matrix of ones

c. The 5×5 identity matrix

d. A 4×4 diagonal matrix, with diagonal entries 3, 4, 2, 5

A useful way to test new ideas in matrix algebra, or to make conjectures, is to make calculations with matrices selected at random. Checking a property for a few matrices does not prove that the property holds in general, but it makes the property more believable. Also, if the property is actually false, making a few calculations may help to discover this.

36. [**M**] Write the command(s) that will create a 5×6 matrix with random entries. In what range of numbers do the entries lie? Tell how to create a 4×4 matrix with random integer entries between -9 and 9. [*Hint:* If x is a random number such that $0 < x < 1$, then $-9.5 < 19(x - .5) < 9.5$.]

37. [**M**] Construct random 4×4 matrices A and B to test whether $AB = BA$. The best way to do this is to compute $AB - BA$ and check whether this difference is the zero matrix. Then test $AB - BA$ for three more pairs of random 4×4 matrices. Report your conclusions.

38. [**M**] Construct a random 5×5 matrix A and test whether $(A + I)(A - I) = A^2 - I$. The best way to do this is to compute $(A + I)(A - I) - (A^2 - I)$ and verify that this difference is the zero matrix. Do this for three random matrices. Then test $(A + B)(A - B) = A^2 - B^2$ the same way for three pairs of random 4×4 matrices. Report your conclusions.

39. [**M**] Use at least three pairs of random 4×4 matrices A and B to test the equalities $(A + B)^T = A^T + B^T$ and $(AB)^T = B^TA^T$, as well as $(AB)^T = A^TB^T$. (See Exercise 37.) Report your conclusions. [*Note*: Most matrix programs use A' for A^T.]

40. [**M**] Let

$$S = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Compute S^k for $k = 2, \dots, 6$.

41. [**M**] Describe in words what happens when A^5, A^{10}, A^{20}, and A^{30} are computed for

$$A = \begin{bmatrix} 1/4 & 1/2 & 1/4 \\ 1/2 & 1/3 & 1/6 \\ 1/4 & 1/6 & 7/12 \end{bmatrix}$$

SOLUTIONS TO PRACTICE PROBLEMS

1. $A\mathbf{x} = \begin{bmatrix} 1 & -3 \\ -2 & 4 \end{bmatrix}\begin{bmatrix} 5 \\ 3 \end{bmatrix} = \begin{bmatrix} -4 \\ 2 \end{bmatrix}$. So $(A\mathbf{x})^T = \begin{bmatrix} -4 & 2 \end{bmatrix}$. Also,

$$\mathbf{x}^TA^T = \begin{bmatrix} 5 & 3 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ -3 & 4 \end{bmatrix} = \begin{bmatrix} -4 & 2 \end{bmatrix}.$$

The quantities $(A\mathbf{x})^T$ and $\mathbf{x}^TA^T$ are equal, by Theorem 3(d). Next,

$$\mathbf{x}\mathbf{x}^T = \begin{bmatrix} 5 \\ 3 \end{bmatrix}\begin{bmatrix} 5 & 3 \end{bmatrix} = \begin{bmatrix} 25 & 15 \\ 15 & 9 \end{bmatrix}$$

$$\mathbf{x}^T\mathbf{x} = \begin{bmatrix} 5 & 3 \end{bmatrix}\begin{bmatrix} 5 \\ 3 \end{bmatrix} = [\,25 + 9\,] = 34$$

A 1×1 matrix such as $\mathbf{x}^T\mathbf{x}$ is usually written without the brackets. Finally, $A^T\mathbf{x}^T$ is not defined, because $\mathbf{x}^T$ does not have two rows to match the two columns of A^T.

2. The fastest way to compute $A^2\mathbf{x}$ is to compute $A(A\mathbf{x})$. The product $A\mathbf{x}$ requires 16 multiplications, 4 for each entry, and $A(A\mathbf{x})$ requires 16 more. In contrast, the product A^2 requires 64 multiplications, 4 for each of the 16 entries in A^2. After that, $A^2\mathbf{x}$ takes 16 more multiplications, for a total of 80.

2.2 THE INVERSE OF A MATRIX

Matrix algebra provides tools for manipulating matrix equations and creating various useful formulas in ways similar to doing ordinary algebra with real numbers. This section investigates the matrix analogue of the reciprocal, or multiplicative inverse, of a nonzero number.

Recall that the multiplicative inverse of a number such as 5 is 1/5 or 5^{-1}. This inverse satisfies the equations

$$5^{-1} \cdot 5 = 1 \quad \text{and} \quad 5 \cdot 5^{-1} = 1$$

The matrix generalization requires *both* equations and avoids the slanted-line notation (for division) because matrix multiplication is not commutative. Furthermore, a full generalization is possible only if the matrices involved are square.[1]

An $n \times n$ matrix A is said to be **invertible** if there is an $n \times n$ matrix C such that

$$CA = I \quad \text{and} \quad AC = I$$

where $I = I_n$, the $n \times n$ identity matrix. In this case, C is an **inverse** of A. In fact, C is uniquely determined by A, because if B were another inverse of A, then $B = BI = B(AC) = (BA)C = IC = C$. This unique inverse is denoted by A^{-1}, so that

$$A^{-1}A = I \quad \text{and} \quad AA^{-1} = I$$

A matrix that is *not* invertible is sometimes called a **singular matrix**, and an invertible matrix is called a **nonsingular matrix**.

EXAMPLE 1 If $A = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}$ and $C = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}$, then

$$AC = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}\begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and}$$

$$CA = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Thus $C = A^{-1}$. ■

Here is a simple formula for the inverse of a 2×2 matrix, along with a test to tell if the inverse exists.

THEOREM 4

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If $ad - bc \neq 0$, then A is invertible and

$$A^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

If $ad - bc = 0$, then A is not invertible.

The simple proof of Theorem 4 is outlined in Exercises 25 and 26. The quantity $ad - bc$ is called the **determinant** of A, and we write

$$\det A = ad - bc$$

Theorem 4 says that a 2×2 matrix A is invertible if and only if $\det A \neq 0$.

[1] One could say that an $m \times n$ matrix A is invertible if there exist $n \times m$ matrices C and D such that $CA = I_n$ and $AD = I_m$. However, these equations imply that A is square and $C = D$. Thus A is invertible as defined above. See Exercises 23–25 in Section 2.1.

EXAMPLE 2 Find the inverse of $A = \begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix}$.

SOLUTION Since $\det A = 3(6) - 4(5) = -2 \neq 0$, A is invertible, and

$$A^{-1} = \frac{1}{-2}\begin{bmatrix} 6 & -4 \\ -5 & 3 \end{bmatrix} = \begin{bmatrix} 6/(-2) & -4/(-2) \\ -5/(-2) & 3/(-2) \end{bmatrix} = \begin{bmatrix} -3 & 2 \\ 5/2 & -3/2 \end{bmatrix}$$

Invertible matrices are indispensable in linear algebra—mainly for algebraic calculations and formula derivations, as in the next theorem. There are also occasions when an inverse matrix provides insight into a mathematical model of a real-life situation, as in Example 3, below.

THEOREM 5

If A is an invertible $n \times n$ matrix, then for each $\mathbf{b}$ in $\mathbb{R}^n$, the equation $A\mathbf{x} = \mathbf{b}$ has the unique solution $\mathbf{x} = A^{-1}\mathbf{b}$.

PROOF Take any $\mathbf{b}$ in $\mathbb{R}^n$. A solution exists because if $A^{-1}\mathbf{b}$ is substituted for $\mathbf{x}$, then $A\mathbf{x} = A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = I\mathbf{b} = \mathbf{b}$. So $A^{-1}\mathbf{b}$ is a solution. To prove that the solution is unique, show that if $\mathbf{u}$ is any solution, then $\mathbf{u}$, in fact, must be $A^{-1}\mathbf{b}$. Indeed, if $A\mathbf{u} = \mathbf{b}$, we can multiply both sides by A^{-1} and obtain

$$A^{-1}A\mathbf{u} = A^{-1}\mathbf{b}, \quad I\mathbf{u} = A^{-1}\mathbf{b}, \quad \text{and} \quad \mathbf{u} = A^{-1}\mathbf{b}$$

EXAMPLE 3 A horizontal elastic beam is supported at each end and is subjected to forces at points 1, 2, 3, as shown in Fig. 1. Let $\mathbf{f}$ in $\mathbb{R}^3$ list the forces at these points, and let $\mathbf{y}$ in $\mathbb{R}^3$ list the amounts of deflection (that is, movement) of the beam at the three points. Using Hooke's law from physics, it can be shown that

$$\mathbf{y} = D\mathbf{f}$$

where D is a *flexibility matrix*. Its inverse is called the *stiffness matrix*. Describe the physical significance of the columns of D and D^{-1}.

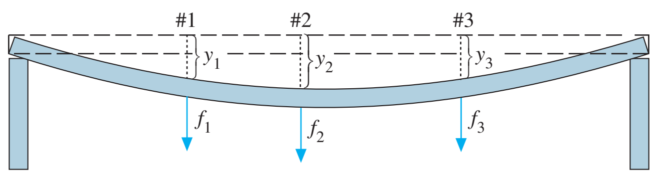


FIGURE 1 Deflection of an elastic beam.

SOLUTION Write $I_3 = [\,\mathbf{e}_1 \quad \mathbf{e}_2 \quad \mathbf{e}_3\,]$ and observe that

$$D = DI_3 = [\,D\mathbf{e}_1 \quad D\mathbf{e}_2 \quad D\mathbf{e}_3\,]$$

Interpret the vector $\mathbf{e}_1 = (1, 0, 0)$ as a unit force applied downward at point 1 on the beam (with zero force at the other two points). Then $D\mathbf{e}_1$, the first column of D, lists the beam deflections due to a unit force at point 1. Similar descriptions apply to the second and third columns of D.

To study the stiffness matrix D^{-1}, observe that the equation $\mathbf{f} = D^{-1}\mathbf{y}$ computes a force vector $\mathbf{f}$ when a deflection vector $\mathbf{y}$ is given. Write

$$D^{-1} = D^{-1}I_3 = [\,D^{-1}\mathbf{e}_1 \quad D^{-1}\mathbf{e}_2 \quad D^{-1}\mathbf{e}_3\,]$$

Now interpret $\mathbf{e}_1$ as a deflection vector. Then $D^{-1}\mathbf{e}_1$ lists the forces that create the deflection. That is, the first column of D^{-1} lists the forces that must be applied at the three points to produce a unit deflection at point 1 and zero deflections at the other points. Similarly, columns 2 and 3 of D^{-1} list the forces required to produce unit deflections at points 2 and 3, respectively. In each column, one or two of the forces must be negative (point upward) to produce a unit deflection at the desired point and zero deflections at the other two points. If the flexibility is measured, for example, in inches of deflection per pound of load, then the stiffness matrix entries are given in pounds of load per inch of deflection. ■

The formula in Theorem 5 is seldom used to solve an equation $A\mathbf{x} = \mathbf{b}$ numerically because row reduction of $[\,A \quad \mathbf{b}\,]$ is nearly always faster. (Row reduction is usually more accurate, too, when computations involve rounding off numbers.) One possible exception is the 2×2 case. In this case, mental computations to solve $A\mathbf{x} = \mathbf{b}$ are sometimes easier using the formula for A^{-1}, as in the next example.

EXAMPLE 4 Use the inverse of the matrix A in Example 2 to solve the system

$$\begin{aligned} 3x_1 + 4x_2 &= 3 \\ 5x_1 + 6x_2 &= 7 \end{aligned}$$

SOLUTION This system is equivalent to $A\mathbf{x} = \mathbf{b}$, so

$$\mathbf{x} = A^{-1}\mathbf{b} = \begin{bmatrix} -3 & 2 \\ 5/2 & -3/2 \end{bmatrix} \begin{bmatrix} 3 \\ 7 \end{bmatrix} = \begin{bmatrix} 5 \\ -3 \end{bmatrix}$$ ■

The next theorem provides three useful facts about invertible matrices.

THEOREM 6

a. If A is an invertible matrix, then A^{-1} is invertible and

$$(A^{-1})^{-1} = A$$

b. If A and B are $n \times n$ invertible matrices, then so is AB, and the inverse of AB is the product of the inverses of A and B in the reverse order. That is,

$$(AB)^{-1} = B^{-1}A^{-1}$$

c. If A is an invertible matrix, then so is A^T, and the inverse of A^T is the transpose of A^{-1}. That is,

$$(A^T)^{-1} = (A^{-1})^T$$

PROOF To verify statement (a), find a matrix C such that

$$A^{-1}C = I \quad \text{and} \quad CA^{-1} = I$$

In fact, these equations are satisfied with A in place of C. Hence A^{-1} is invertible, and A is its inverse. Next, to prove statement (b), compute:

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$$

A similar calculation shows that $(B^{-1}A^{-1})(AB) = I$. For statement (c), use Theorem 3(d), read from right to left, $(A^{-1})^T A^T = (AA^{-1})^T = I^T = I$. Similarly, $A^T(A^{-1})^T = I^T = I$. Hence A^T is invertible, and its inverse is $(A^{-1})^T$. ■

The following generalization of Theorem 6(b) is needed later.

The product of $n \times n$ invertible matrices is invertible, and the inverse is the product of their inverses in the reverse order.

There is an important connection between invertible matrices and row operations that leads to a method for computing inverses. As we shall see, an invertible matrix A is row equivalent to an identity matrix, and we can find A^{-1} by *watching the row reduction of A to I.*

Elementary Matrices

An **elementary matrix** is one that is obtained by performing a single elementary row operation on an identity matrix. The next example illustrates the three kinds of elementary matrices.

EXAMPLE 5 Let

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{bmatrix},$$

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

Compute E_1A, E_2A, and E_3A, and describe how these products can be obtained by elementary row operations on A.

SOLUTION Verify that

$$E_1A = \begin{bmatrix} a & b & c \\ d & e & f \\ g-4a & h-4b & i-4c \end{bmatrix}, \quad E_2A = \begin{bmatrix} d & e & f \\ a & b & c \\ g & h & i \end{bmatrix},$$

$$E_3A = \begin{bmatrix} a & b & c \\ d & e & f \\ 5g & 5h & 5i \end{bmatrix}.$$

Addition of -4 times row 1 of A to row 3 produces E_1A. (This is a row replacement operation.) An interchange of rows 1 and 2 of A produces E_2A, and multiplication of row 3 of A by 5 produces E_3A. ■

Left-multiplication (that is, multiplication on the left) by E_1 in Example 5 has the same effect on any $3 \times n$ matrix. It adds -4 times row 1 to row 3. In particular, since $E_1 \cdot I = E_1$, we see that E_1 *itself* is produced by this same row operation on the identity. Thus Example 5 illustrates the following general fact about elementary matrices. See Exercises 27 and 28.

> If an elementary row operation is performed on an $m \times n$ matrix A, the resulting matrix can be written as EA, where the $m \times m$ matrix E is created by performing the same row operation on I_m.

Since row operations are reversible, as shown in Section 1.1, elementary matrices are invertible, for if E is produced by a row operation on I, then there is another row operation of the same type that changes E back into I. Hence there is an elementary matrix F such that $FE = I$. Since E and F correspond to reverse operations, $EF = I$, too.

Each elementary matrix E is invertible. The inverse of E is the elementary matrix of the same type that transforms E back into I.

EXAMPLE 6 Find the inverse of $E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}$.

SOLUTION To transform E_1 into I, add $+4$ times row 1 to row 3. The elementary matrix that does this is

$$E_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ +4 & 0 & 1 \end{bmatrix}$$

The following theorem provides the best way to "visualize" an invertible matrix, and the theorem leads immediately to a method for finding the inverse of a matrix.

THEOREM 7

An $n \times n$ matrix A is invertible if and only if A is row equivalent to I_n, and in this case, any sequence of elementary row operations that reduces A to I_n also transforms I_n into A^{-1}.

PROOF Suppose that A is invertible. Then, since the equation $A\mathbf{x} = \mathbf{b}$ has a solution for each $\mathbf{b}$ (Theorem 5), A has a pivot position in every row (Theorem 4 in Section 1.4). Because A is square, the n pivot positions must be on the diagonal, which implies that the reduced echelon form of A is I_n. That is, $A \sim I_n$.

Now suppose, conversely, that $A \sim I_n$. Then, since each step of the row reduction of A corresponds to left-multiplication by an elementary matrix, there exist elementary matrices $E_1, \dots, E_p$ such that

$$A \sim E_1 A \sim E_2(E_1 A) \sim \cdots \sim E_p(E_{p-1} \cdots E_1 A) = I_n$$

That is,

$$E_p \cdots E_1 A = I_n \tag{1}$$

Since the product $E_p \cdots E_1$ of invertible matrices is invertible, (1) leads to

$$\begin{aligned} (E_p \cdots E_1)^{-1}(E_p \cdots E_1)A &= (E_p \cdots E_1)^{-1} I_n \\ A &= (E_p \cdots E_1)^{-1} \end{aligned}$$

Thus A is invertible, as it is the inverse of an invertible matrix (Theorem 6). Also,

$$A^{-1} = [\,(E_p \cdots E_1)^{-1}\,]^{-1} = E_p \cdots E_1$$

Then $A^{-1} = E_p \cdots E_1 \cdot I_n$, which says that A^{-1} results from applying $E_1, \dots, E_p$ successively to I_n. This is the same sequence in (1) that reduced A to I_n.

An Algorithm for Finding A^{-1}

If we place A and I side-by-side to form an augmented matrix $[\,A \quad I\,]$, then row operations on this matrix produce identical operations on A and on I. By Theorem 7, either there are row operations that transform A to I_n and I_n to A^{-1} or else A is not invertible.

ALGORITHM FOR FINDING A^{-1}

Row reduce the augmented matrix $[\,A \quad I\,]$. If A is row equivalent to I, then $[\,A \quad I\,]$ is row equivalent to $[\,I \quad A^{-1}\,]$. Otherwise, A does not have an inverse.

EXAMPLE 7 Find the inverse of the matrix $A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix}$, if it exists.

SOLUTION

$$
\begin{aligned}
[\,A \quad I\,] &= \begin{bmatrix} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{bmatrix} \\
&\sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & -3 & -4 & 0 & -4 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 3 & -4 & 1 \end{bmatrix} \\
&\sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3/2 & -2 & 1/2 \end{bmatrix} \\
&\sim \begin{bmatrix} 1 & 0 & 0 & -9/2 & 7 & -3/2 \\ 0 & 1 & 0 & -2 & 4 & -1 \\ 0 & 0 & 1 & 3/2 & -2 & 1/2 \end{bmatrix}
\end{aligned}
$$

Theorem 7 shows, since $A \sim I$, that A is invertible, and

$$
A^{-1} = \begin{bmatrix} -9/2 & 7 & -3/2 \\ -2 & 4 & -1 \\ 3/2 & -2 & 1/2 \end{bmatrix}
$$

It is a good idea to check the final answer:

$$
AA^{-1} = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix} \begin{bmatrix} -9/2 & 7 & -3/2 \\ -2 & 4 & -1 \\ 3/2 & -2 & 1/2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}
$$

It is not necessary to check that $A^{-1}A = I$ since A is invertible. ■

Another View of Matrix Inversion

Denote the columns of I_n by $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Then row reduction of $[\,A \quad I\,]$ to $[\,I \quad A^{-1}\,]$ can be viewed as the simultaneous solution of the n systems

$$
A\mathbf{x} = \mathbf{e}_1, \quad A\mathbf{x} = \mathbf{e}_2, \quad \ldots, \quad A\mathbf{x} = \mathbf{e}_n \tag{2}
$$

where the "augmented columns" of these systems have all been placed next to A to form $[\,A \quad \mathbf{e}_1 \quad \mathbf{e}_2 \quad \cdots \quad \mathbf{e}_n\,] = [\,A \quad I\,]$. The equation $AA^{-1} = I$ and the definition of matrix multiplication show that the columns of A^{-1} are precisely the solutions of the systems in (2). This observation is useful because some applied problems may require finding only one or two columns of A^{-1}. In this case, only the corresponding systems in (2) need be solved.

NUMERICAL NOTE

WEB

In practical work, A^{-1} is seldom computed, unless the entries of A^{-1} are needed. Computing both A^{-1} and $A^{-1}\mathbf{b}$ takes about three times as many arithmetic operations as solving $A\mathbf{x} = \mathbf{b}$ by row reduction, and row reduction may be more accurate.

PRACTICE PROBLEMS

1. Use determinants to determine which of the following matrices are invertible.

a. $\begin{bmatrix} 3 & -9 \\ 2 & 6 \end{bmatrix}$ b. $\begin{bmatrix} 4 & -9 \\ 0 & 5 \end{bmatrix}$ c. $\begin{bmatrix} 6 & -9 \\ -4 & 6 \end{bmatrix}$

2. Find the inverse of the matrix $A = \begin{bmatrix} 1 & -2 & -1 \\ -1 & 5 & 6 \\ 5 & -4 & 5 \end{bmatrix}$, if it exists.

2.2 EXERCISES

Find the inverses of the matrices in Exercises 1–4.

1. $\begin{bmatrix} 8 & 6 \\ 5 & 4 \end{bmatrix}$ **2.** $\begin{bmatrix} 3 & 2 \\ 8 & 5 \end{bmatrix}$

3. $\begin{bmatrix} 7 & 3 \\ -6 & -3 \end{bmatrix}$ **4.** $\begin{bmatrix} 2 & -4 \\ 4 & -6 \end{bmatrix}$

5. Use the inverse found in Exercise 1 to solve the system

$$\begin{aligned} 8x_1 + 6x_2 &= 2 \\ 5x_1 + 4x_2 &= -1 \end{aligned}$$

6. Use the inverse found in Exercise 3 to solve the system

$$\begin{aligned} 7x_1 + 3x_2 &= -9 \\ -6x_1 - 3x_2 &= 4 \end{aligned}$$

7. Let $A = \begin{bmatrix} 1 & 2 \\ 5 & 12 \end{bmatrix}$, $\mathbf{b}_1 = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 1 \\ -5 \end{bmatrix}$, $\mathbf{b}_3 = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$, and $\mathbf{b}_4 = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$.

a. Find A^{-1}, and use it to solve the four equations

$$A\mathbf{x} = \mathbf{b}_1, \quad A\mathbf{x} = \mathbf{b}_2, \quad A\mathbf{x} = \mathbf{b}_3, \quad A\mathbf{x} = \mathbf{b}_4$$

b. The four equations in part (a) can be solved by the *same* set of row operations, since the coefficient matrix is the same in each case. Solve the four equations in part (a) by row reducing the augmented matrix $[\, A \;\; \mathbf{b}_1 \;\; \mathbf{b}_2 \;\; \mathbf{b}_3 \;\; \mathbf{b}_4 \,]$.

8. Suppose P is invertible and $A = PBP^{-1}$. Solve for B in terms of A.

In Exercises 9 and 10, mark each statement True or False. Justify each answer.

9. a. In order for a matrix B to be the inverse of A, the equations $AB = I$ and $BA = I$ must both be true.

b. If A and B are $n \times n$ and invertible, then $A^{-1}B^{-1}$ is the inverse of AB.

c. If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $ab - cd \neq 0$, then A is invertible.

d. If A is an invertible $n \times n$ matrix, then the equation $A\mathbf{x} = \mathbf{b}$ is consistent for *each* $\mathbf{b}$ in $\mathbb{R}^n$.

e. Each elementary matrix is invertible.

10. a. If A is invertible, then elementary row operations that reduce A to the identity I_n also reduce A^{-1} to I_n.

b. If A is invertible, then the inverse of A^{-1} is A itself.

c. A product of invertible $n \times n$ matrices is invertible, and the inverse of the product is the product of their inverses in the same order.

d. If A is an $n \times n$ matrix and $A\mathbf{x} = \mathbf{e}_j$ is consistent for every $j \in \{1, 2, \ldots, n\}$, then A is invertible. Note: $\mathbf{e}_1, \ldots, \mathbf{e}_n$ represent the columns of the identity matrix.

e. If A can be row reduced to the identity matrix, then A must be invertible.

11. Let A be an invertible $n \times n$ matrix, and let B be an $n \times p$ matrix. Show that the equation $AX = B$ has a unique solution $A^{-1}B$.

12. Use matrix algebra to show that if A is invertible and D satisfies $AD = I$, then $D = A^{-1}$.

13. Suppose $AB = AC$, where B and C are $n \times p$ matrices and A is invertible. Show that $B = C$. Is this true, in general, when A is not invertible?

14. Suppose $(B - C)D = 0$, where B and C are $m \times n$ matrices and D is invertible. Show that $B = C$.

15. Let A be an invertible $n \times n$ matrix, and let B be an $n \times p$ matrix. Explain why $A^{-1}B$ can be computed by row reduction:

If $[\,A \quad B\,] \sim \cdots \sim [\,I \quad X\,]$, then $X = A^{-1}B$.

If A is larger than 2×2, then row reduction of $[A \quad B]$ is much faster than computing both A^{-1} and $A^{-1}B$.

16. Suppose A and B are $n \times n$ matrices, B is invertible, and AB is invertible. Show that A is invertible. [*Hint:* Let $C = AB$, and solve this equation for A.]

17. Suppose A, B, and C are invertible $n \times n$ matrices. Show that ABC is also invertible by producing a matrix D such that $(ABC)D = I$ and $D(ABC) = I$.

18. Solve the equation $AB = BC$ for A, assuming that A, B, and C are square and B is invertible.

19. If A, B, and C are $n \times n$ invertible matrices, does the equation $C^{-1}(A + X)B^{-1} = I_n$ have a solution, X? If so, find it.

20. Suppose A, B, and X are $n \times n$ matrices with A, X, and $A - AX$ invertible, and suppose

$$(A - AX)^{-1} = X^{-1}B \tag{3}$$

a. Explain why B is invertible.

b. Solve equation (3) for X. If a matrix needs to be inverted, explain why that matrix is invertible.

21. Explain why the columns of an $n \times n$ matrix A are linearly independent when A is invertible.

22. Explain why the columns of an $n \times n$ matrix A span $\mathbb{R}^n$ when A is invertible. [*Hint:* Review Theorem 4 in Section 1.4.]

23. Suppose A is $n \times n$ and the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution. Explain why A has n pivot columns and A is row equivalent to I_n. By Theorem 7, this shows that A must be invertible. (This exercise and Exercise 24 will be cited in Section 2.3.)

24. Suppose A is $n \times n$ and the equation $A\mathbf{x} = \mathbf{b}$ has a solution for each $\mathbf{b}$ in $\mathbb{R}^n$. Explain why A must be invertible. [*Hint:* Is A row equivalent to I_n?]

Exercises 25 and 26 prove Theorem 4 for $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

25. Show that if $ad - bc = 0$, then the equation $A\mathbf{x} = \mathbf{0}$ has more than one solution. Why does this imply that A is not invertible? [*Hint:* First, consider $a = b = 0$. Then, if a and b are not both zero, consider the vector $\mathbf{x} = \begin{bmatrix} -b \\ a \end{bmatrix}$.]

26. Show that if $ad - bc \neq 0$, the formula for A^{-1} works.

Exercises 27 and 28 prove special cases of the facts about elementary matrices stated in the box following Example 5. Here A is a 3×3 matrix and $I = I_3$. (A general proof would require slightly more notation.)

27. Let A be a 3×3 matrix.

a. Use equation (2) from Section 2.1 to show that $\text{row}_i(A) = \text{row}_i(I) \cdot A$, for $i = 1, 2, 3$.

b. Show that if rows 1 and 2 of A are interchanged, then the result may be written as EA, where E is an elementary matrix formed by interchanging rows 1 and 2 of I.

c. Show that if row 3 of A is multiplied by 5, then the result may be written as EA, where E is formed by multiplying row 3 of I by 5.

28. Suppose row 2 of A is replaced by $\text{row}_2(A) - 3 \cdot \text{row}_1(A)$. Show that the result is EA, where E is formed from I by replacing $\text{row}_2(I)$ by $\text{row}_2(I) - 3 \cdot \text{row}_1(A)$.

Find the inverses of the matrices in Exercises 29–32, if they exist. Use the algorithm introduced in this section.

29. $\begin{bmatrix} 1 & -3 \\ 4 & -9 \end{bmatrix}$

30. $\begin{bmatrix} 3 & 6 \\ 4 & 7 \end{bmatrix}$

31. $\begin{bmatrix} 1 & 0 & -2 \\ -3 & 1 & 4 \\ 2 & -3 & 4 \end{bmatrix}$

32. $\begin{bmatrix} 1 & 2 & -1 \\ -4 & -7 & 3 \\ -2 & -6 & 4 \end{bmatrix}$

33. Use the algorithm from this section to find the inverses of

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$

Let A be the corresponding $n \times n$ matrix, and let B be its inverse. Guess the form of B, and then show that $AB = I$.

34. Repeat the strategy of Exercise 33 to guess the inverse B of

$$A = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 2 & 2 & 0 & & 0 \\ 3 & 3 & 3 & & 0 \\ \vdots & & & \ddots & \vdots \\ n & n & n & \cdots & n \end{bmatrix}.$$

Show that $AB = I$.

35. Let $A = \begin{bmatrix} -1 & -7 & -3 \\ 2 & 15 & 6 \\ 1 & 3 & 2 \end{bmatrix}$. Find the third column of A^{-1} without computing the other columns.

36. **[M]** Let $A = \begin{bmatrix} -25 & -9 & -27 \\ 536 & 185 & 537 \\ 154 & 52 & 143 \end{bmatrix}$. Find the second and third columns of A^{-1} without computing the first column.

37. Let $A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 1 & 5 \end{bmatrix}$. Construct a 2×3 matrix C (by trial and error) using only 1, -1, and 0 as entries, such that $CA = I_2$. Compute AC and note that $AC \neq I_3$.

38. Let $A = \begin{bmatrix} 1 & -1 & 1 & 0 \\ 0 & 1 & -1 & 1 \end{bmatrix}$. Construct a 4×2 matrix D using only 1 and 0 as entries, such that $AD = I_2$. Is it possible that $CA = I_4$ for some 4×2 matrix C? Why or why not?

39. [**M**] Let

$$D = \begin{bmatrix} .011 & .003 & .001 \\ .003 & .009 & .003 \\ .001 & .003 & .011 \end{bmatrix}$$

be a flexibility matrix, with flexibility measured in inches per pound. Suppose that forces of 40, 50, and 30 lb are applied at points 1, 2, and 3, respectively, in Fig. 1 of Example 3. Find the corresponding deflections.

40. [**M**] Compute the stiffness matrix D^{-1} for D in Exercise 39. List the forces needed to produce a deflection of .04 in. at point 3, with zero deflections at the other points.

41. [**M**] Let

$$D = \begin{bmatrix} .0130 & .0050 & .0020 & .0010 \\ .0050 & .0100 & .0040 & .0020 \\ .0020 & .0040 & .0100 & .0050 \\ .0010 & .0020 & .0050 & .0130 \end{bmatrix}$$

be a flexibility matrix for an elastic beam such as the one in Example 3, with four points at which force is applied. Units are centimeters per newton of force. Measurements at the four points show deflections of .07, .12, .16, and .12 cm. Determine the forces at the four points.

42. [**M**] With D as in Exercise 41, determine the forces that produce a deflection of .22 cm at the second point on the beam, with zero deflections at the other three points. How is the answer related to the entries in D^{-1}? [*Hint:* First answer the question when the deflection is 1 cm at the second point.]

SOLUTIONS TO PRACTICE PROBLEMS

1. a. $\det \begin{bmatrix} 3 & -9 \\ 2 & 6 \end{bmatrix} = 3 \cdot 6 - (-9) \cdot 2 = 18 + 18 = 36$. The determinant is nonzero, so the matrix is invertible.

b. $\det \begin{bmatrix} 4 & -9 \\ 0 & 5 \end{bmatrix} = 4 \cdot 5 - (-9) \cdot 0 = 20 \neq 0$. The matrix is invertible.

c. $\det \begin{bmatrix} 6 & -9 \\ -4 & 6 \end{bmatrix} = 6 \cdot 6 - (-9)(-4) = 36 - 36 = 0$. The matrix is not invertible.

2.
$$[\,A \quad I\,] \sim \begin{bmatrix} 1 & -2 & -1 & 1 & 0 & 0 \\ -1 & 5 & 6 & 0 & 1 & 0 \\ 5 & -4 & 5 & 0 & 0 & 1 \end{bmatrix}$$
$$\sim \begin{bmatrix} 1 & -2 & -1 & 1 & 0 & 0 \\ 0 & 3 & 5 & 1 & 1 & 0 \\ 0 & 6 & 10 & -5 & 0 & 1 \end{bmatrix}$$
$$\sim \begin{bmatrix} 1 & -2 & -1 & 1 & 0 & 0 \\ 0 & 3 & 5 & 1 & 1 & 0 \\ 0 & 0 & 0 & -7 & -2 & 1 \end{bmatrix}$$

So $[\,A \quad I\,]$ is row equivalent to a matrix of the form $[\,B \quad D\,]$, where B is square and has a row of zeros. Further row operations will not transform B into I, so we stop. A does not have an inverse.

2.3 CHARACTERIZATIONS OF INVERTIBLE MATRICES

This section provides a review of most of the concepts introduced in Chapter 1, in relation to systems of n linear equations in n unknowns and to *square* matrices. The main result is Theorem 8.

THEOREM 8

The Invertible Matrix Theorem

Let A be a square $n \times n$ matrix. Then the following statements are equivalent. That is, for a given A, the statements are either all true or all false.

a. A is an invertible matrix.
b. A is row equivalent to the $n \times n$ identity matrix.
c. A has n pivot positions.
d. The equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
e. The columns of A form a linearly independent set.
f. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.
g. The equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$.
h. The columns of A span $\mathbb{R}^n$.
i. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps $\mathbb{R}^n$ onto $\mathbb{R}^n$.
j. There is an $n \times n$ matrix C such that $CA = I$.
k. There is an $n \times n$ matrix D such that $AD = I$.
l. A^T is an invertible matrix.

$$\begin{array}{ccccc} & & \text{(a)} & & \\ & \nearrow & & \searrow & \\ \text{(b)} & & & & \text{(j)} \\ \Uparrow & & & & \Downarrow \\ \text{(c)} & & \Leftarrow & & \text{(d)} \end{array}$$

FIGURE 1

First, we need some notation. If the truth of statement (a) always implies that statement (j) is true, we say that (a) *implies* (j) and write (a) $\Rightarrow$ (j). The proof will establish the "circle" of implications shown in Fig. 1. If any one of these five statements is true, then so are the others. Finally, the proof will link the remaining statements of the theorem to the statements in this circle.

PROOF If statement (a) is true, then A^{-1} works for C in (j), so (a) $\Rightarrow$ (j). Next, (j) $\Rightarrow$ (d) by Exercise 23 in Section 2.1. (Turn back and read the exercise.) Also, (d) $\Rightarrow$ (c) by Exercise 23 in Section 2.2. If A is square and has n pivot positions, then the pivots must lie on the main diagonal, in which case the reduced echelon form of A is I_n. Thus (c) $\Rightarrow$ (b). Also, (b) $\Rightarrow$ (a) by Theorem 7 in Section 2.2. This completes the circle in Fig. 1.

$$\begin{array}{ccc} & \text{(k)} & \\ \nearrow & & \searrow \\ \text{(a)} & \Leftarrow & \text{(g)} \end{array}$$

$$\text{(g)} \Leftrightarrow \text{(h)} \Leftrightarrow \text{(i)}$$

$$\text{(d)} \Leftrightarrow \text{(e)} \Leftrightarrow \text{(f)}$$

$$\text{(a)} \Leftrightarrow \text{(l)}$$

Next, (a) $\Rightarrow$ (k) because A^{-1} works for D. Also, (k) $\Rightarrow$ (g) by Exercise 26 in Section 2.1, and (g) $\Rightarrow$ (a) by Exercise 24 in Section 2.2. So (k) and (g) are linked to the circle. Further, (g), (h), and (i) are equivalent for any matrix, by Theorem 4 in Section 1.4 and Theorem 12(a) in Section 1.9. Thus, (h) and (i) are linked through (g) to the circle.

Since (d) is linked to the circle, so are (e) and (f), because (d), (e), and (f) are all equivalent for *any* matrix A. (See Section 1.7 and Theorem 12(b) in Section 1.9.) Finally, (a) $\Rightarrow$ (l) by Theorem 6(c) in Section 2.2, and (l) $\Rightarrow$ (a) by the same theorem with A and A^T interchanged. This completes the proof. ■

Because of Theorem 5 in Section 2.2, statement (g) in Theorem 8 could also be written as "The equation $A\mathbf{x} = \mathbf{b}$ has a *unique* solution for each $\mathbf{b}$ in $\mathbb{R}^n$." This statement certainly implies (b) and hence implies that A is invertible.

The next fact follows from Theorem 8 and Exercise 12 in Section 2.2.

Let A and B be square matrices. If $AB = I$, then A and B are both invertible, with $B = A^{-1}$ and $A = B^{-1}$.

The Invertible Matrix Theorem divides the set of all $n \times n$ matrices into two disjoint classes: the invertible (nonsingular) matrices, and the noninvertible (singular) matrices. Each statement in the theorem describes a property of every $n \times n$ invertible matrix. The *negation* of a statement in the theorem describes a property of every $n \times n$ singular matrix. For instance, an $n \times n$ singular matrix is *not* row equivalent to I_n, does *not* have n pivot positions, and has linearly *dependent* columns. Negations of other statements are considered in the exercises.

EXAMPLE 1 Use the Invertible Matrix Theorem to decide if A is invertible:

$$A = \begin{bmatrix} 1 & 0 & -2 \\ 3 & 1 & -2 \\ -5 & -1 & 9 \end{bmatrix}$$

SOLUTION

$$A \sim \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 4 \\ 0 & -1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 4 \\ 0 & 0 & 3 \end{bmatrix}$$

So A has three pivot positions and hence is invertible, by the Invertible Matrix Theorem, statement (c). ■

SG Expanded Table for the IMT 2–10

The power of the Invertible Matrix Theorem lies in the connections it provides among so many important concepts, such as linear independence of columns of a matrix A and the existence of solutions to equations of the form $A\mathbf{x} = \mathbf{b}$. It should be emphasized, however, that the Invertible Matrix Theorem *applies only to square matrices*. For example, if the columns of a 4×3 matrix are linearly independent, we cannot use the Invertible Matrix Theorem to conclude anything about the existence or nonexistence of solutions to equations of the form $A\mathbf{x} = \mathbf{b}$.

Invertible Linear Transformations

Recall from Section 2.1 that matrix multiplication corresponds to composition of linear transformations. When a matrix A is invertible, the equation $A^{-1}A\mathbf{x} = \mathbf{x}$ can be viewed as a statement about linear transformations. See Fig. 2.

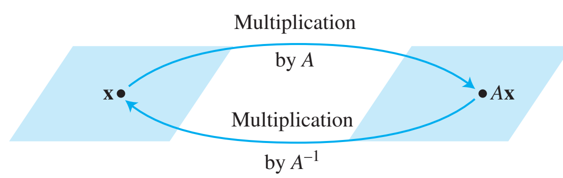


FIGURE 2 A^{-1} transforms $A\mathbf{x}$ back to $\mathbf{x}$.

A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is said to be **invertible** if there exists a function $S : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$S(T(\mathbf{x})) = \mathbf{x} \quad \text{for all } \mathbf{x} \text{ in } \mathbb{R}^n \tag{1}$$

$$T(S(\mathbf{x})) = \mathbf{x} \quad \text{for all } \mathbf{x} \text{ in } \mathbb{R}^n \tag{2}$$

The next theorem shows that if such an S exists, it is unique and must be a linear transformation. We call S the **inverse** of T and write it as T^{-1}.

THEOREM 9

Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation and let A be the standard matrix for T. Then T is invertible if and only if A is an invertible matrix. In that case, the linear transformation S given by $S(\mathbf{x}) = A^{-1}\mathbf{x}$ is the unique function satisfying equations (1) and (2).

PROOF Suppose that T is invertible. Then (2) shows that T is onto $\mathbb{R}^n$, for if $\mathbf{b}$ is in $\mathbb{R}^n$ and $\mathbf{x} = S(\mathbf{b})$, then $T(\mathbf{x}) = T(S(\mathbf{b})) = \mathbf{b}$, so each $\mathbf{b}$ is in the range of T. Thus A is invertible, by the Invertible Matrix Theorem, statement (i).

Conversely, suppose that A is invertible, and let $S(\mathbf{x}) = A^{-1}\mathbf{x}$. Then, S is a linear transformation, and S obviously satisfies (1) and (2). For instance,

$$S(T(\mathbf{x})) = S(A\mathbf{x}) = A^{-1}(A\mathbf{x}) = \mathbf{x}$$

Thus T is invertible. The proof that S is unique is outlined in Exercise 38. ■

EXAMPLE 2 What can you say about a one-to-one linear transformation T from $\mathbb{R}^n$ into $\mathbb{R}^n$?

SOLUTION The columns of the standard matrix A of T are linearly independent (by Theorem 12 in Section 1.9). So A is invertible, by the Invertible Matrix Theorem, and T maps $\mathbb{R}^n$ *onto* $\mathbb{R}^n$. Also, T is invertible, by Theorem 9. ■

NUMERICAL NOTES

In practical work, you might occasionally encounter a "nearly singular" or **ill-conditioned** matrix—an invertible matrix that can become singular if some of its entries are changed ever so slightly. In this case, row reduction may produce fewer than n pivot positions, as a result of roundoff error. Also, roundoff error can sometimes make a singular matrix appear to be invertible.

WEB

Some matrix programs will compute a **condition number** for a square matrix. The larger the condition number, the closer the matrix is to being singular. The condition number of the identity matrix is 1. A singular matrix has an infinite condition number. In extreme cases, a matrix program may not be able to distinguish between a singular matrix and an ill-conditioned matrix.

Exercises 41–45 show that matrix computations can produce substantial error when a condition number is large.

PRACTICE PROBLEMS

1. Determine if $A = \begin{bmatrix} 2 & 3 & 4 \\ 2 & 3 & 4 \\ 2 & 3 & 4 \end{bmatrix}$ is invertible.
2. Suppose that for a certain $n \times n$ matrix A, statement (g) of the Invertible Matrix Theorem is *not* true. What can you say about equations of the form $A\mathbf{x} = \mathbf{b}$?
3. Suppose that A and B are $n \times n$ matrices and the equation $AB\mathbf{x} = \mathbf{0}$ has a nontrivial solution. What can you say about the matrix AB?

2.3 EXERCISES

Unless otherwise specified, assume that all matrices in these exercises are $n \times n$. Determine which of the matrices in Exercises 1–10 are invertible. Use as few calculations as possible. Justify your answers.

1. $\begin{bmatrix} 5 & 7 \\ -3 & -6 \end{bmatrix}$

2. $\begin{bmatrix} -4 & 2 \\ 6 & -3 \end{bmatrix}$

3. $\begin{bmatrix} 3 & 0 & 0 \\ -3 & -4 & 0 \\ 8 & 5 & -3 \end{bmatrix}$

4. $\begin{bmatrix} -5 & 1 & 4 \\ 0 & 0 & 0 \\ 1 & 4 & 9 \end{bmatrix}$

5. $\begin{bmatrix} 3 & 0 & -3 \\ 2 & 0 & 4 \\ -4 & 0 & 7 \end{bmatrix}$

6. $\begin{bmatrix} 1 & -3 & -6 \\ 0 & 4 & 3 \\ -3 & 6 & 0 \end{bmatrix}$

7. $\begin{bmatrix} -1 & -3 & 0 & 1 \\ 3 & 5 & 8 & -3 \\ -2 & -6 & 3 & 2 \\ 0 & -1 & 2 & 1 \end{bmatrix}$

8. $\begin{bmatrix} 3 & 4 & 7 & 4 \\ 0 & 1 & 4 & 6 \\ 0 & 0 & 2 & 8 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

9. [**M**] $\begin{bmatrix} 4 & 0 & -3 & -7 \\ -6 & 9 & 9 & 9 \\ 7 & -5 & 10 & 19 \\ -1 & 2 & 4 & -1 \end{bmatrix}$

10. [**M**] $\begin{bmatrix} 5 & 3 & 1 & 7 & 9 \\ 6 & 4 & 2 & 8 & -8 \\ 7 & 5 & 3 & 10 & 9 \\ 9 & 6 & 4 & -9 & -5 \\ 8 & 5 & 2 & 11 & 4 \end{bmatrix}$

In Exercises 11 and 12, the matrices are all $n \times n$. Each part of the exercises is an *implication* of the form "If $\langle$ statement 1 $\rangle$, then $\langle$ statement 2 $\rangle$." Mark an implication as True if the truth of $\langle$ statement 2 $\rangle$ *always* follows whenever $\langle$ statement 1 $\rangle$ happens to be true. An implication is False if there is an instance in which $\langle$ statement 2 $\rangle$ is false but $\langle$ statement 1 $\rangle$ is true. Justify each answer.

11. a. If the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution, then A is row equivalent to the $n \times n$ identity matrix.

b. If the columns of A span $\mathbb{R}^n$, then the columns are linearly independent.

c. If A is an $n \times n$ matrix, then the equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$.

d. If the equation $A\mathbf{x} = \mathbf{0}$ has a nontrivial solution, then A has fewer than n pivot positions.

e. If A^T is not invertible, then A is not invertible.

12. a. If there is an $n \times n$ matrix D such that $AD = I$, then $DA = I$.

b. If the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps $\mathbb{R}^n$ into $\mathbb{R}^n$, then the row reduced echelon form of A is I.

c. If the columns of A are linearly independent, then the columns of A span $\mathbb{R}^n$.

d. If the equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$, then the transformation $\mathbf{x} \mapsto A\mathbf{x}$ is not one-to-one.

e. If there is a $\mathbf{b}$ in $\mathbb{R}^n$ such that the equation $A\mathbf{x} = \mathbf{b}$ is consistent, then the solution is unique.

13. An $m \times n$ **upper triangular matrix** is one whose entries *below* the main diagonal are 0's (as in Exercise 8). When is a square upper triangular matrix invertible? Justify your answer.

14. An $m \times n$ **lower triangular matrix** is one whose entries *above* the main diagonal are 0's (as in Exercise 3). When is a square lower triangular matrix invertible? Justify your answer.

15. Is it possible for a 4×4 matrix to be invertible when its columns do not span $\mathbb{R}^4$? Why or why not?

16. If an $n \times n$ matrix A is invertible, then the columns of A^T are linearly independent. Explain why.

17. Can a square matrix with two identical columns be invertible? Why or why not?

18. Can a square matrix with two identical rows be invertible? Why or why not?

19. If the columns of a 7×7 matrix D are linearly independent, what can be said about the solutions of $D\mathbf{x} = \mathbf{b}$? Why?

20. If A is a 5×5 matrix and the equation $A\mathbf{x} = \mathbf{b}$ is consistent for every $\mathbf{b}$ in $\mathbb{R}^5$, is it possible that for some $\mathbf{b}$, the equation $A\mathbf{x} = \mathbf{b}$ has more than one solution? Why or why not?

21. If the equation $C\mathbf{u} = \mathbf{v}$ has more than one solution for some $\mathbf{v}$ in $\mathbb{R}^n$, can the columns of the $n \times n$ matrix C span $\mathbb{R}^n$? Why or why not?

22. If $n \times n$ matrices E and F have the property that $EF = I$, then E and F commute. Explain why.

23. Assume that F is an $n \times n$ matrix. If the equation $F\mathbf{x} = \mathbf{y}$ is inconsistent for some $\mathbf{y}$ in $\mathbb{R}^n$, what can you say about the equation $F\mathbf{x} = \mathbf{0}$? Why?

24. If an $n \times n$ matrix G cannot be row reduced to I_n, what can you say about the columns of G? Why?

25. Verify the boxed statement preceding Example 1.

26. Explain why the columns of A^2 span $\mathbb{R}^n$ whenever the columns of an $n \times n$ matrix A are linearly independent.

27. Let A and B be $n \times n$ matrices. Show that if AB is invertible, so is A. You cannot use Theorem 6(b), because you cannot *assume* that A and B are invertible. [*Hint:* There is a matrix W such that $ABW = I$. Why?]

28. Let A and B be $n \times n$ matrices. Show that if AB is invertible, so is B.

29. If A is an $n \times n$ matrix and the transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one, what else can you say about this transformation? Justify your answer.

30. If A is an $n \times n$ matrix and the equation $A\mathbf{x} = \mathbf{b}$ has more than one solution for some $\mathbf{b}$, then the transformation $\mathbf{x} \mapsto A\mathbf{x}$ is not one-to-one. What else can you say about this transformation? Justify your answer.

31. Suppose A is an $n \times n$ matrix with the property that the equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$. Without using Theorems 5 or 8, explain why each equation $A\mathbf{x} = \mathbf{b}$ has in fact exactly one solution.

32. Suppose A is an $n \times n$ matrix with the property that the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution. Without using the Invertible Matrix Theorem, explain directly why the equation $A\mathbf{x} = \mathbf{b}$ must have a solution for each $\mathbf{b}$ in $\mathbb{R}^n$.

In Exercises 33 and 34, T is a linear transformation from $\mathbb{R}^2$ into $\mathbb{R}^2$. Show that T is invertible and find a formula for T^{-1}.

33. $T(x_1, x_2) = (-5x_1 + 9x_2, 4x_1 - 7x_2)$

34. $T(x_1, x_2) = (2x_1 - 8x_2, -2x_1 + 7x_2)$

35. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be an invertible linear transformation. Explain why T is both one-to-one and onto $\mathbb{R}^n$. Use equations (1) and (2). Then give a second explanation using one or more theorems.

36. Suppose a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ has the property that $T(\mathbf{u}) = T(\mathbf{v})$ for some pair of distinct vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$. Can T map $\mathbb{R}^n$ onto $\mathbb{R}^n$? Why or why not?

37. Suppose T and U are linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^n$ such that $T(U(\mathbf{x})) = \mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$. Is it true that $U(T(\mathbf{x})) = \mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$? Why or why not?

38. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be an invertible linear transformation, and let S and U be functions from $\mathbb{R}^n$ into $\mathbb{R}^n$ such that $S(T(\mathbf{x})) = \mathbf{x}$ and $U(T(\mathbf{x})) = \mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$. Show that $U(\mathbf{v}) = S(\mathbf{v})$ for all $\mathbf{v}$ in $\mathbb{R}^n$. This will show that T has a unique inverse, as asserted in Theorem 9. [*Hint:* Given any $\mathbf{v}$ in $\mathbb{R}^n$, we can write $\mathbf{v} = T(\mathbf{x})$ for some $\mathbf{x}$. Why? Compute $S(\mathbf{v})$ and $U(\mathbf{v})$.]

39. Let T be a linear transformation that maps $\mathbb{R}^n$ onto $\mathbb{R}^n$. Show that T^{-1} exists and maps $\mathbb{R}^n$ onto $\mathbb{R}^n$. Is T^{-1} also one-to-one?

40. Suppose T and S satisfy the invertibility equations (1) and (2), where T is a linear transformation. Show directly that S is a linear transformation. [*Hint:* Given $\mathbf{u}, \mathbf{v}$ in $\mathbb{R}^n$, let $\mathbf{x} = S(\mathbf{u})$, $\mathbf{y} = S(\mathbf{v})$. Then $T(\mathbf{x}) = \mathbf{u}$, $T(\mathbf{y}) = \mathbf{v}$. Why? Apply S to both sides of the equation $T(\mathbf{x}) + T(\mathbf{y}) = T(\mathbf{x} + \mathbf{y})$. Also, consider $T(c\mathbf{x}) = cT(\mathbf{x})$.]

41. **[M]** Suppose an experiment leads to the following system of equations:

$$\begin{aligned} 4.5x_1 + 3.1x_2 &= 19.249 \\ 1.6x_1 + 1.1x_2 &= 6.843 \end{aligned} \tag{3}$$

a. Solve system (3), and then solve system (4), below, in which the data on the right have been rounded to two decimal places. In each case, find the *exact* solution.

$$\begin{aligned} 4.5x_1 + 3.1x_2 &= 19.25 \\ 1.6x_1 + 1.1x_2 &= 6.84 \end{aligned} \tag{4}$$

b. The entries in system (4) differ from those in system (3) by less than .05%. Find the percentage error when using the solution of (4) as an approximation for the solution of (3).

c. Use a matrix program to produce the condition number of the coefficient matrix in (3).

Exercises 42–44 show how to use the condition number of a matrix A to estimate the accuracy of a computed solution of $A\mathbf{x} = \mathbf{b}$. If the entries of A and $\mathbf{b}$ are accurate to about r significant digits and if the condition number of A is approximately 10^k (with k a positive integer), then the computed solution of $A\mathbf{x} = \mathbf{b}$ should usually be accurate to at least $r - k$ significant digits.

42. **[M]** Let A be the matrix in Exercise 9. Find the condition number of A. Construct a random vector $\mathbf{x}$ in $\mathbb{R}^4$ and compute $\mathbf{b} = A\mathbf{x}$. Then use a matrix program to compute the solution $\mathbf{x}_1$ of $A\mathbf{x} = \mathbf{b}$. To how many digits do $\mathbf{x}$ and $\mathbf{x}_1$ agree? Find out the number of digits the matrix program stores accurately, and report how many digits of accuracy are lost when $\mathbf{x}_1$ is used in place of the exact solution $\mathbf{x}$.

43. **[M]** Repeat Exercise 42 for the matrix in Exercise 10.

44. **[M]** Solve an equation $A\mathbf{x} = \mathbf{b}$ for a suitable $\mathbf{b}$ to find the last column of the inverse of the *fifth-order Hilbert matrix*

$$A = \begin{bmatrix} 1 & 1/2 & 1/3 & 1/4 & 1/5 \\ 1/2 & 1/3 & 1/4 & 1/5 & 1/6 \\ 1/3 & 1/4 & 1/5 & 1/6 & 1/7 \\ 1/4 & 1/5 & 1/6 & 1/7 & 1/8 \\ 1/5 & 1/6 & 1/7 & 1/8 & 1/9 \end{bmatrix}$$

How many digits in each entry of $\mathbf{x}$ do you expect to be correct? Explain. [*Note:* The exact solution is (630, −12600, 56700, −88200, 44100).]

45. **[M]** Some matrix programs, such as MATLAB, have a command to create Hilbert matrices of various sizes. If possible, use an inverse command to compute the inverse of a twelfth-order or larger Hilbert matrix, A. Compute AA^{-1}. Report what you find.

SG Mastering: Reviewing and Reflecting 2–13

SOLUTIONS TO PRACTICE PROBLEMS

1. The columns of A are obviously linearly dependent because columns 2 and 3 are multiples of column 1. Hence A cannot be invertible, by the Invertible Matrix Theorem.
2. If statement (g) is *not* true, then the equation $A\mathbf{x} = \mathbf{b}$ is inconsistent for at least one $\mathbf{b}$ in $\mathbb{R}^n$.
3. Apply the Invertible Matrix Theorem to the matrix AB in place of A. Then statement (d) becomes: $AB\mathbf{x} = \mathbf{0}$ has only the trivial solution. This is not true. So AB is not invertible.

2.4 PARTITIONED MATRICES

A key feature of our work with matrices has been the ability to regard a matrix A as a list of column vectors rather than just a rectangular array of numbers. This point of view has been so useful that we wish to consider other **partitions** of A, indicated by horizontal and vertical dividing rules, as in Example 1 below. Partitioned matrices appear in most modern applications of linear algebra because the notation highlights essential structures in matrix analysis, as in the chapter introductory example on aircraft design. This section provides an opportunity to review matrix algebra and use the Invertible Matrix Theorem.

EXAMPLE 1 The matrix

$$A = \left[\begin{array}{rrr|rr|r} 3 & 0 & -1 & 5 & 9 & -2 \\ -5 & 2 & 4 & 0 & -3 & 1 \\ \hline -8 & -6 & 3 & 1 & 7 & -4 \end{array}\right]$$

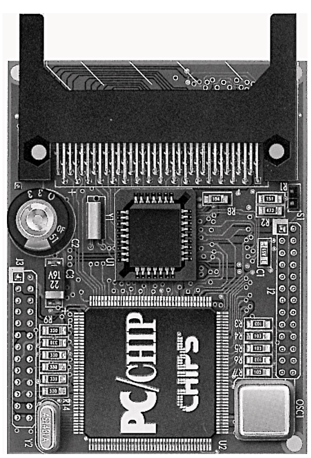
can also be written as the 2×3 **partitioned** (or **block**) **matrix**

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix}$$

whose entries are the *blocks* (or *submatrices*)

$$A_{11} = \begin{bmatrix} 3 & 0 & -1 \\ -5 & 2 & 4 \end{bmatrix}, \quad A_{12} = \begin{bmatrix} 5 & 9 \\ 0 & -3 \end{bmatrix}, \quad A_{13} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

$$A_{21} = \begin{bmatrix} -8 & -6 & 3 \end{bmatrix}, \quad A_{22} = \begin{bmatrix} 1 & 7 \end{bmatrix}, \quad A_{23} = \begin{bmatrix} -4 \end{bmatrix}$$ ■



EXAMPLE 2 When a matrix A appears in a mathematical model of a physical system such as an electrical network, a transportation system, or a large corporation, it may be natural to regard A as a partitioned matrix. For instance, if a microcomputer circuit board consists mainly of three VLSI (very large-scale integrated) microchips, then the matrix for the circuit board might have the general form

$$A = \left[\begin{array}{c|c|c} A_{11} & A_{12} & A_{13} \\ \hline A_{21} & A_{22} & A_{23} \\ \hline A_{31} & A_{32} & A_{33} \end{array}\right]$$

The submatrices on the "diagonal" of A—namely, A_{11}, A_{22}, and A_{33}—concern the three VLSI chips, while the other submatrices depend on the interconnections among those microchips. ■

For the induction step, assume A and B are $(k+1) \times (k+1)$ matrices, and partition A and B in a form similar to that displayed in Exercise 23.

25. Without using row reduction, find the inverse of

$$A = \begin{bmatrix} 1 & 2 & 0 & 0 & 0 \\ 3 & 5 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 7 & 8 \\ 0 & 0 & 0 & 5 & 6 \end{bmatrix}$$

26. [**M**] For block operations, it may be necessary to access or enter submatrices of a large matrix. Describe the functions or commands of a matrix program that accomplish the following tasks. Suppose A is a 20×30 matrix.

a. Display the submatrix of A from rows 5 to 10 and columns 15 to 20.

b. Insert a 5×10 matrix B into A, beginning at row 5 and column 10.

c. Create a 50×50 matrix of the form $C = \begin{bmatrix} A & 0 \\ 0 & A^T \end{bmatrix}$.

[*Note:* It may not be necessary to specify the zero blocks in C.]

27. [**M**] Suppose memory or size restrictions prevent a matrix program from working with matrices having more than 32 rows and 32 columns, and suppose some project involves 50×50 matrices A and B. Describe the commands or operations of the matrix program that accomplish the following tasks.

a. Compute $A + B$.

b. Compute AB.

c. Solve $A\mathbf{x} = \mathbf{b}$ for some vector $\mathbf{b}$ in $\mathbb{R}^{50}$, assuming that A can be partitioned into a 2×2 block matrix $[A_{ij}]$, with A_{11} an invertible 20×20 matrix, A_{22} an invertible 30×30 matrix, and A_{12} a zero matrix. [*Hint:* Describe appropriate smaller systems to solve, without using any matrix inverses.]

SOLUTIONS TO PRACTICE PROBLEMS

1. If $\begin{bmatrix} I & 0 \\ A & I \end{bmatrix}$ is invertible, its inverse has the form $\begin{bmatrix} W & X \\ Y & Z \end{bmatrix}$. Verify that

$$\begin{bmatrix} I & 0 \\ A & I \end{bmatrix}\begin{bmatrix} W & X \\ Y & Z \end{bmatrix} = \begin{bmatrix} W & X \\ AW + Y & AX + Z \end{bmatrix}$$

So W, X, Y, Z must satisfy $W = I$, $X = 0$, $AW + Y = 0$, and $AX + Z = I$. It follows that $Y = -A$ and $Z = I$. Hence

$$\begin{bmatrix} I & 0 \\ A & I \end{bmatrix}\begin{bmatrix} I & 0 \\ -A & I \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

The product in the reverse order is also the identity, so the block matrix is invertible, and its inverse is $\begin{bmatrix} I & 0 \\ -A & I \end{bmatrix}$. (You could also appeal to the Invertible Matrix Theorem.)

2. $X^TX = \begin{bmatrix} X_1^T \\ X_2^T \end{bmatrix}\begin{bmatrix} X_1 & X_2 \end{bmatrix} = \begin{bmatrix} X_1^TX_1 & X_1^TX_2 \\ X_2^TX_1 & X_2^TX_2 \end{bmatrix}$. The partitions of X^T and X are automatically conformable for block multiplication because the columns of X^T are the rows of X. This partition of X^TX is used in several computer algorithms for matrix computations.

2.5 MATRIX FACTORIZATIONS

A *factorization* of a matrix A is an equation that expresses A as a product of two or more matrices. Whereas matrix multiplication involves a *synthesis* of data (combining the effects of two or more linear transformations into a single matrix), matrix factorization is an *analysis* of data. In the language of computer science, the expression of A as a product amounts to a *preprocessing* of the data in A, organizing that data into two or more parts whose structures are more useful in some way, perhaps more accessible for computation.

Matrix factorizations and, later, factorizations of linear transformations will appear at a number of key points throughout the text. This section focuses on a factorization that lies at the heart of several important computer programs widely used in applications, such as the airflow problem described in the chapter introduction. Several other factorizations, to be studied later, are introduced in the exercises.

The LU Factorization

The LU factorization, described below, is motivated by the fairly common industrial and business problem of solving a sequence of equations, all with the same coefficient matrix:

$$A\mathbf{x} = \mathbf{b}_1, \quad A\mathbf{x} = \mathbf{b}_2, \quad \ldots, \quad A\mathbf{x} = \mathbf{b}_p \tag{1}$$

See Exercise 32, for example. Also see Section 5.8, where the inverse power method is used to estimate eigenvalues of a matrix by solving equations like those in sequence (1), one at a time.

When A is invertible, one could compute A^{-1} and then compute $A^{-1}\mathbf{b}_1$, $A^{-1}\mathbf{b}_2$, and so on. However, it is more efficient to solve the first equation in sequence (1) by row reduction and obtain an LU factorization of A at the same time. Thereafter, the remaining equations in sequence (1) are solved with the LU factorization.

At first, assume that A is an $m \times n$ matrix that can be row reduced to echelon form, *without row interchanges*. (Later, we will treat the general case.) Then A can be written in the form $A = LU$, where L is an $m \times m$ lower triangular matrix with 1's on the diagonal and U is an $m \times n$ echelon form of A. For instance, see Fig. 1. Such a factorization is called an **LU factorization** of A. The matrix L is invertible and is called a *unit* lower triangular matrix.

$$A = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} \blacksquare & * & * & * & * \\ 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}}_{U}$$

FIGURE 1 An LU factorization.

Before studying how to construct L and U, we should look at why they are so useful. When $A = LU$, the equation $A\mathbf{x} = \mathbf{b}$ can be written as $L(U\mathbf{x}) = \mathbf{b}$. Writing $\mathbf{y}$ for $U\mathbf{x}$, we can find $\mathbf{x}$ by solving the *pair* of equations

$$\boxed{\begin{aligned} L\mathbf{y} &= \mathbf{b} \\ U\mathbf{x} &= \mathbf{y} \end{aligned}} \tag{2}$$

First solve $L\mathbf{y} = \mathbf{b}$ for $\mathbf{y}$, and then solve $U\mathbf{x} = \mathbf{y}$ for $\mathbf{x}$. See Fig. 2. Each equation is easy to solve because L and U are triangular.

EXAMPLE 1 It can be verified that

$$A = \begin{bmatrix} 3 & -7 & -2 & 2 \\ -3 & 5 & 1 & 0 \\ 6 & -4 & 0 & -5 \\ -9 & 5 & -5 & 12 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 2 & -5 & 1 & 0 \\ -3 & 8 & 3 & 1 \end{bmatrix} \begin{bmatrix} 3 & -7 & -2 & 2 \\ 0 & -2 & -1 & 2 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix} = LU$$

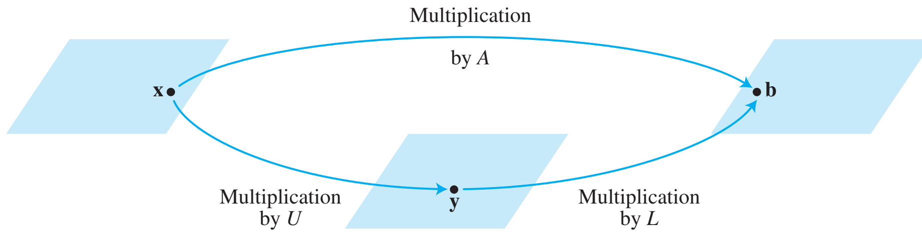


FIGURE 2 Factorization of the mapping $\mathbf{x} \mapsto A\mathbf{x}$.

Use this LU factorization of A to solve $A\mathbf{x} = \mathbf{b}$, where $\mathbf{b} = \begin{bmatrix} -9 \\ 5 \\ 7 \\ 11 \end{bmatrix}$.

SOLUTION The solution of $L\mathbf{y} = \mathbf{b}$ needs only 6 multiplications and 6 additions, because the arithmetic takes place only in column 5. (The zeros below each pivot in L are created automatically by the choice of row operations.)

$$\begin{bmatrix} L & \mathbf{b} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & -9 \\ -1 & 1 & 0 & 0 & 5 \\ 2 & -5 & 1 & 0 & 7 \\ -3 & 8 & 3 & 1 & 11 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 & -9 \\ 0 & 1 & 0 & 0 & -4 \\ 0 & 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} I & \mathbf{y} \end{bmatrix}$$

Then, for $U\mathbf{x} = \mathbf{y}$, the "backward" phase of row reduction requires 4 divisions, 6 multiplications, and 6 additions. (For instance, creating the zeros in column 4 of $[\,U \quad \mathbf{y}\,]$ requires 1 division in row 4 and 3 multiplication–addition pairs to add multiples of row 4 to the rows above.)

$$\begin{bmatrix} U & \mathbf{y} \end{bmatrix} = \begin{bmatrix} 3 & -7 & -2 & 2 & -9 \\ 0 & -2 & -1 & 2 & -4 \\ 0 & 0 & -1 & 1 & 5 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 & 3 \\ 0 & 1 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 & -6 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} 3 \\ 4 \\ -6 \\ -1 \end{bmatrix}$$

To find $\mathbf{x}$ requires 28 arithmetic operations, or "flops" (floating point operations), excluding the cost of finding L and U. In contrast, row reduction of $[\,A \quad \mathbf{b}\,]$ to $[\,I \quad \mathbf{x}\,]$ takes 62 operations. ■

The computational efficiency of the LU factorization depends on knowing L and U. The next algorithm shows that the row reduction of A to an echelon form U amounts to an LU factorization because it produces L with essentially no extra work. After the first row reduction, L and U are available for solving additional equations whose coefficient matrix is A.

An LU Factorization Algorithm

Suppose A can be reduced to an echelon form U using only row replacements that add a multiple of one row to another row *below it*. In this case, there exist unit lower triangular elementary matrices $E_1, \dots, E_p$ such that

$$E_p \cdots E_1 A = U \tag{3}$$

Then

$$A = (E_p \cdots E_1)^{-1} U = LU$$

where

$$L = (E_p \cdots E_1)^{-1} \tag{4}$$

It can be shown that products and inverses of unit lower triangular matrices are also unit lower triangular. (For instance, see Exercise 19.) Thus L is unit lower triangular.

Note that the row operations in equation (3), which reduce A to U, also reduce the L in equation (4) to I, because $E_p \cdots E_1 L = (E_p \cdots E_1)(E_p \cdots E_1)^{-1} = I$. This observation is the key to *constructing* L.

ALGORITHM FOR AN LU FACTORIZATION

1. Reduce A to an echelon form U by a sequence of row replacement operations, if possible.
2. Place entries in L such that the *same sequence of row operations* reduces L to I.

Step 1 is not always possible, but when it is, the argument above shows that an LU factorization exists. Example 2 will show how to implement step 2. By construction, L will satisfy

$$(E_p \cdots E_1)L = I$$

using the same $E_1, \dots, E_p$ as in equation (3). Thus L will be invertible, by the Invertible Matrix Theorem, with $(E_p \cdots E_1) = L^{-1}$. From (3), $L^{-1}A = U$, and $A = LU$. So step 2 will produce an acceptable L.

EXAMPLE 2 Find an LU factorization of

$$A = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ -6 & 0 & 7 & -3 & 1 \end{bmatrix}$$

SOLUTION Since A has four rows, L should be 4×4. The first column of L is the first column of A divided by the top pivot entry:

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & & 1 & 0 \\ -3 & & & 1 \end{bmatrix}$$

Compare the first columns of A and L. *The row operations that create zeros in the first column of A will also create zeros in the first column of L.* To make this same correspondence of row operations on A hold for the rest of L, watch a row reduction of A to an echelon form U. That is, *highlight the entries* in each matrix that are used to determine the sequence of row operations that transform A into U. [See the highlighted entries in equation (5).]

$$A = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ -6 & 0 & 7 & -3 & 1 \end{bmatrix} \sim \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & -9 & -3 & -4 & 10 \\ 0 & 12 & 4 & 12 & -5 \end{bmatrix} = A_1 \tag{5}$$

$$\sim A_2 = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 4 & 7 \end{bmatrix} \sim \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix} = U$$

The highlighted entries above determine the row reduction of A to U. At each pivot column, divide the highlighted entries by the pivot and place the result into L:

$$\begin{bmatrix} 2 \\ -4 \\ 2 \\ -6 \end{bmatrix} \begin{bmatrix} 3 \\ -9 \\ 12 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} \begin{bmatrix} 5 \end{bmatrix}$$

$$\div 2 \qquad \div 3 \qquad \div 2 \qquad \div 5$$

$$\downarrow \qquad \downarrow \qquad \downarrow \qquad \downarrow$$

$$\begin{bmatrix} 1 & & & \\ -2 & 1 & & \\ 1 & -3 & 1 & \\ -3 & 4 & 2 & 1 \end{bmatrix}, \quad \text{and} \quad L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & -3 & 1 & 0 \\ -3 & 4 & 2 & 1 \end{bmatrix}$$

An easy calculation verifies that this L and U satisfy $LU = A$. ■

In practical work, row interchanges are nearly always needed, because partial pivoting is used for high accuracy. (Recall that this procedure selects, among the possible choices for a pivot, an entry in the column having the largest absolute value.) To handle row interchanges, the LU factorization above can be modified easily to produce an L that is *permuted lower triangular*, in the sense that a rearrangement (called a permutation) of the rows of L can make L (unit) lower triangular. The resulting *permuted LU factorization* solves $A\mathbf{x} = \mathbf{b}$ in the same way as before, except that the reduction of $[\,L \quad \mathbf{b}\,]$ to $[\,I \quad \mathbf{y}\,]$ follows the order of the pivots in L from left to right, starting with the pivot in the first column. A reference to an "LU factorization" usually includes the possibility that L might be permuted lower triangular. For details, see the *Study Guide*.

SG Permuted LU Factorizations 2-23

NUMERICAL NOTES

The following operation counts apply to an $n \times n$ dense matrix A (with most entries nonzero) for n moderately large, say, $n \geq 30$.[1]

1. Computing an LU factorization of A takes about $2n^3/3$ flops (about the same as row reducing $[\,A \quad \mathbf{b}\,]$), whereas finding A^{-1} requires about $2n^3$ flops.
2. Solving $L\mathbf{y} = \mathbf{b}$ and $U\mathbf{x} = \mathbf{y}$ requires about $2n^2$ flops, because any $n \times n$ triangular system can be solved in about n^2 flops.
3. Multiplication of $\mathbf{b}$ by A^{-1} also requires about $2n^2$ flops, but the result may not be as accurate as that obtained from L and U (because of roundoff error when computing both A^{-1} and $A^{-1}\mathbf{b}$).
4. If A is sparse (with mostly zero entries), then L and U may be sparse, too, whereas A^{-1} is likely to be dense. In this case, a solution of $A\mathbf{x} = \mathbf{b}$ with an LU factorization is *much* faster than using A^{-1}. See Exercise 31.

WEB

A Matrix Factorization in Electrical Engineering

Matrix factorization is intimately related to the problem of constructing an electrical network with specified properties. The following discussion gives just a glimpse of the connection between factorization and circuit design.

---

[1] See Section 3.8 in *Applied Linear Algebra*, 3rd ed., by Ben Noble and James W. Daniel (Englewood Cliffs, NJ: Prentice-Hall, 1988). Recall that for our purposes, a *flop* is $+$, $-$, $\times$, or $\div$.

Suppose the box in Fig. 3 represents some sort of electric circuit, with an input and output. Record the input voltage and current by $\begin{bmatrix} v_1 \\ i_1 \end{bmatrix}$ (with voltage v in volts and current i in amps), and record the output voltage and current by $\begin{bmatrix} v_2 \\ i_2 \end{bmatrix}$. Frequently, the transformation $\begin{bmatrix} v_1 \\ i_1 \end{bmatrix} \mapsto \begin{bmatrix} v_2 \\ i_2 \end{bmatrix}$ is linear. That is, there is a matrix A, called the *transfer matrix*, such that

$$\begin{bmatrix} v_2 \\ i_2 \end{bmatrix} = A \begin{bmatrix} v_1 \\ i_1 \end{bmatrix}$$

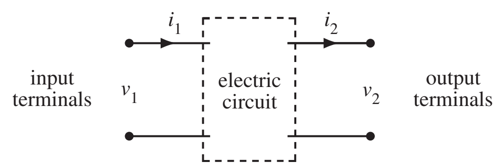


FIGURE 3 A circuit with input and output terminals.

Figure 4 shows a *ladder network*, where two circuits (there could be more) are connected in series, so that the output of one circuit becomes the input of the next circuit. The left circuit in Fig. 4 is called a *series circuit*, with resistance R_1 (in ohms).

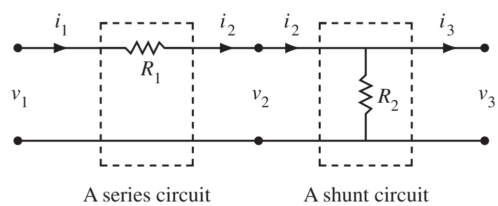


FIGURE 4 A ladder network.

The right circuit in Fig. 4 is a *shunt circuit*, with resistance R_2. Using Ohm's law and Kirchhoff's laws, one can show that the transfer matrices of the series and shunt circuits, respectively, are

$$\underbrace{\begin{bmatrix} 1 & -R_1 \\ 0 & 1 \end{bmatrix}}_{\text{Transfer matrix of series circuit}} \quad \text{and} \quad \underbrace{\begin{bmatrix} 1 & 0 \\ -1/R_2 & 1 \end{bmatrix}}_{\text{Transfer matrix of shunt circuit}}$$

EXAMPLE 3

a. Compute the transfer matrix of the ladder network in Fig. 4.

b. Design a ladder network whose transfer matrix is $\begin{bmatrix} 1 & -8 \\ -.5 & 5 \end{bmatrix}$.

SOLUTION

a. Let A_1 and A_2 be the transfer matrices of the series and shunt circuits, respectively. Then an input vector $\mathbf{x}$ is transformed first into $A_1\mathbf{x}$ and then into $A_2(A_1\mathbf{x})$. The series connection of the circuits corresponds to composition of linear transformations, and the transfer matrix of the ladder network is (note the order)

$$A_2A_1 = \begin{bmatrix} 1 & 0 \\ -1/R_2 & 1 \end{bmatrix}\begin{bmatrix} 1 & -R_1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -R_1 \\ -1/R_2 & 1 + R_1/R_2 \end{bmatrix} \tag{6}$$

b. To factor the matrix $\begin{bmatrix} 1 & -8 \\ -.5 & 5 \end{bmatrix}$ into the product of transfer matrices, as in equation (6), look for R_1 and R_2 in Fig. 4 to satisfy

$$\begin{bmatrix} 1 & -R_1 \\ -1/R_2 & 1 + R_1/R_2 \end{bmatrix} = \begin{bmatrix} 1 & -8 \\ -.5 & 5 \end{bmatrix}$$

From the $(1, 2)$-entries, $R_1 = 8$ ohms, and from the $(2, 1)$-entries, $1/R_2 = .5$ ohm and $R_2 = 1/.5 = 2$ ohms. With these values, the network in Fig. 4 has the desired transfer matrix. ■

A network transfer matrix summarizes the input–output behavior (the design specifications) of the network without reference to the interior circuits. To physically build a network with specified properties, an engineer first determines if such a network can be constructed (or *realized*). Then the engineer tries to factor the transfer matrix into matrices corresponding to smaller circuits that perhaps are already manufactured and ready for assembly. In the common case of alternating current, the entries in the transfer matrix are usually rational complex-valued functions. (See Exercises 19 and 20 in Section 2.4 and Example 2 in Section 3.3.) A standard problem is to find a *minimal realization* that uses the smallest number of electrical components.

PRACTICE PROBLEM

Find an LU factorization of $A = \begin{bmatrix} 2 & -4 & -2 & 3 \\ 6 & -9 & -5 & 8 \\ 2 & -7 & -3 & 9 \\ 4 & -2 & -2 & -1 \\ -6 & 3 & 3 & 4 \end{bmatrix}$. [*Note*: It will turn out that A has only three pivot columns, so the method of Example 2 will produce only the first three columns of L. The remaining two columns of L come from I_5.]

2.5 EXERCISES

In Exercises 1–6, solve the equation $A\mathbf{x} = \mathbf{b}$ by using the LU factorization given for A. In Exercises 1 and 2, also solve $A\mathbf{x} = \mathbf{b}$ by ordinary row reduction.

1. $A = \begin{bmatrix} 3 & -7 & -2 \\ -3 & 5 & 1 \\ 6 & -4 & 0 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -7 \\ 5 \\ 2 \end{bmatrix}$

$A = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & -5 & 1 \end{bmatrix}\begin{bmatrix} 3 & -7 & -2 \\ 0 & -2 & -1 \\ 0 & 0 & -1 \end{bmatrix}$

2. $A = \begin{bmatrix} 2 & -6 & 4 \\ -4 & 8 & 0 \\ 0 & -4 & 6 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 2 \\ -4 \\ 6 \end{bmatrix}$

$A = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} 2 & -6 & 4 \\ 0 & -4 & 8 \\ 0 & 0 & -2 \end{bmatrix}$

3. $A = \begin{bmatrix} 2 & -4 & 2 \\ -4 & 5 & 2 \\ 6 & -9 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 6 \\ 0 \\ 6 \end{bmatrix}$

$A = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 3 & -1 & 1 \end{bmatrix}\begin{bmatrix} 2 & -4 & 2 \\ 0 & -3 & 6 \\ 0 & 0 & 1 \end{bmatrix}$

4. $A = \begin{bmatrix} 1 & -1 & 2 \\ 1 & -3 & 1 \\ 3 & 7 & 5 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 0 \\ -5 \\ 7 \end{bmatrix}$

$A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 3 & -5 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 & 2 \\ 0 & -2 & -1 \\ 0 & 0 & -6 \end{bmatrix}$

5. $A = \begin{bmatrix} 1 & -2 & -2 & -3 \\ 3 & -9 & 0 & -9 \\ -1 & 2 & 4 & 7 \\ -3 & -6 & 26 & 2 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 6 \\ 0 \\ 3 \end{bmatrix}$

$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -3 & 4 & -2 & 1 \end{bmatrix}\begin{bmatrix} 1 & -2 & -2 & -3 \\ 0 & -3 & 6 & 0 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

6. $A = \begin{bmatrix} 1 & 3 & 2 & 0 \\ -2 & -3 & -4 & 12 \\ 3 & 0 & 4 & -36 \\ -5 & -3 & -8 & 49 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 1 \\ -2 \\ -1 \\ 2 \end{bmatrix}$

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 3 & -3 & 1 & 0 \\ -5 & 4 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 2 & 0 \\ 0 & 3 & 0 & 12 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Find an LU factorization of the matrices in Exercises 7–16 (with L unit lower triangular). Note that MATLAB will usually produce a permuted LU factorization because it uses partial pivoting for numerical accuracy.

7. $\begin{bmatrix} 2 & 5 \\ -3 & -4 \end{bmatrix}$

8. $\begin{bmatrix} 6 & 4 \\ 12 & 5 \end{bmatrix}$

9. $\begin{bmatrix} 3 & 1 & 2 \\ -9 & 0 & -4 \\ 9 & 9 & 14 \end{bmatrix}$

10. $\begin{bmatrix} -5 & 0 & 4 \\ 10 & 2 & -5 \\ 10 & 10 & 16 \end{bmatrix}$

11. $\begin{bmatrix} 3 & 7 & 2 \\ 6 & 19 & 4 \\ -3 & -2 & 3 \end{bmatrix}$

12. $\begin{bmatrix} 2 & 3 & 2 \\ 4 & 13 & 9 \\ -6 & 5 & 4 \end{bmatrix}$

13. $\begin{bmatrix} 1 & 3 & -5 & -3 \\ -1 & -5 & 8 & 4 \\ 4 & 2 & -5 & -7 \\ -2 & -4 & 7 & 5 \end{bmatrix}$

14. $\begin{bmatrix} 1 & 3 & 1 & 5 \\ 5 & 20 & 6 & 31 \\ -2 & -1 & -1 & -4 \\ -1 & 7 & 1 & 7 \end{bmatrix}$

15. $\begin{bmatrix} 2 & 0 & 5 & 2 \\ -6 & 3 & -13 & -3 \\ 4 & 9 & 16 & 17 \end{bmatrix}$

16. $\begin{bmatrix} 2 & -3 & 4 \\ -4 & 8 & -7 \\ 6 & -5 & 14 \\ -6 & 9 & -12 \\ 8 & -6 & 19 \end{bmatrix}$

17. When A is invertible, MATLAB finds A^{-1} by factoring $A = LU$ (where L may be permuted lower triangular), inverting L and U, and then computing $U^{-1}L^{-1}$. Use this method to compute the inverse of A in Exercise 2. (Apply the algorithm in Section 2.2 to L and to U.)

18. Find A^{-1} as in Exercise 17, using A from Exercise 3.

19. Let A be a lower triangular $n \times n$ matrix with nonzero entries on the diagonal. Show that A is invertible and A^{-1} is lower triangular. [*Hint:* Explain why A can be changed into I using only row replacements and scaling. (Where are the pivots?) Also, explain why the row operations that reduce A to I change I into a lower triangular matrix.]

20. Let $A = LU$ be an LU factorization. Explain why A can be row reduced to U using only replacement operations. (This fact is the converse of what was proved in the text.)

21. Suppose $A = BC$, where B is invertible. Show that any sequence of row operations that reduces B to I also reduces A to C. The converse is not true, since the zero matrix may be factored as $0 = B \cdot 0$.

Exercises 22–26 provide a glimpse of some widely used matrix factorizations, some of which are discussed later in the text.

22. (*Reduced LU Factorization*) With A as in the Practice Problem, find a 5×3 matrix B and a 3×4 matrix C such that $A = BC$. Generalize this idea to the case where A is $m \times n$, $A = LU$, and U has only three nonzero rows.

23. (*Rank Factorization*) Suppose an $m \times n$ matrix A admits a factorization $A = CD$ where C is $m \times 4$ and D is $4 \times n$.

a. Show that A is the sum of four outer products. (See Section 2.4.)

b. Let $m = 400$ and $n = 100$. Explain why a computer programmer might prefer to store the data from A in the form of two matrices C and D.

24. (*QR Factorization*) Suppose $A = QR$, where Q and R are $n \times n$, R is invertible and upper triangular, and Q has the property that $Q^TQ = I$. Show that for each $\mathbf{b}$ in $\mathbb{R}^n$, the equation $A\mathbf{x} = \mathbf{b}$ has a unique solution. What computations with Q and R will produce the solution?

WEB

25. (*Singular Value Decomposition*) Suppose $A = UDV^T$, where U and V are $n \times n$ matrices with the property that $U^TU = I$ and $V^TV = I$, and where D is a diagonal matrix with positive numbers $\sigma_1, \dots, \sigma_n$ on the diagonal. Show that A is invertible, and find a formula for A^{-1}.

26. (*Spectral Factorization*) Suppose a 3×3 matrix A admits a factorization as $A = PDP^{-1}$, where P is some invertible 3×3 matrix and D is the diagonal matrix

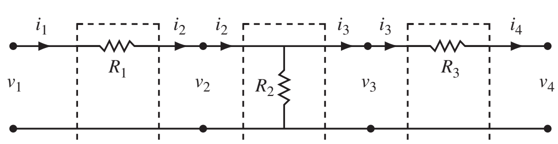
$$D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Show that this factorization is useful when computing high powers of A. Find fairly simple formulas for A^2, A^3, and A^k (k a positive integer), using P and the entries in D.

27. Design two different ladder networks that each output 9 volts and 4 amps when the input is 12 volts and 6 amps.

28. Show that if three shunt circuits (with resistances R_1, R_2, R_3) are connected in series, the resulting network has the same transfer matrix as a single shunt circuit. Find a formula for the resistance in that circuit.

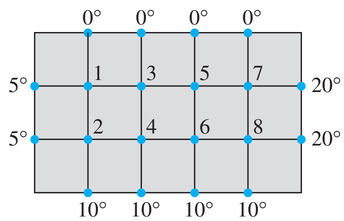
29. a. Compute the transfer matrix of the network in the figure below.



b. Let $A = \begin{bmatrix} 3 & -12 \\ -1/3 & 5/3 \end{bmatrix}$. Design a ladder network whose transfer matrix is A by finding a suitable matrix factorization of A.

30. Find a different factorization of the transfer matrix A in Exercise 29, and thereby design a different ladder network whose transfer matrix is A.

31. [**M**] Consider the heat plate in the following figure (refer to Exercise 33 in Section 1.1).



The solution to the steady-state heat flow problem for this plate is approximated by the solution to the equation $A\mathbf{x} = \mathbf{b}$, where $\mathbf{b} = (5, 15, 0, 10, 0, 10, 20, 30)$ and

$$A = \begin{bmatrix} 4 & -1 & -1 & & & & & \\ -1 & 4 & 0 & -1 & & & & \\ -1 & 0 & 4 & -1 & -1 & & & \\ & -1 & -1 & 4 & 0 & -1 & & \\ & & -1 & 0 & 4 & -1 & -1 & \\ & & & -1 & -1 & 4 & 0 & -1 \\ & & & & -1 & 0 & 4 & -1 \\ & & & & & -1 & -1 & 4 \end{bmatrix}$$

WEB

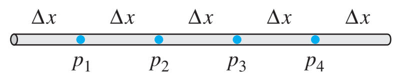
The missing entries in A are zeros. The nonzero entries of A lie within a band along the main diagonal. Such *band matrices* occur in a variety of applications and often are extremely large (with thousands of rows and columns but relatively narrow bands).

a. Use the method in Example 2 to construct an LU factorization of A, and note that both factors are band matrices (with two nonzero diagonals below or above the main diagonal). Compute $LU - A$ to check your work.

b. Use the LU factorization to solve $A\mathbf{x} = \mathbf{b}$.

c. Obtain A^{-1} and note that A^{-1} is a dense matrix with no band structure. When A is large, L and U can be stored in much less space than A^{-1}. This fact is another reason for preferring the LU factorization of A to A^{-1} itself.

32. [**M**] The band matrix A shown below can be used to estimate the unsteady conduction of heat in a rod when the temperatures at points $p_1, \dots, p_4$ on the rod change with time.[2]



The constant C in the matrix depends on the physical nature of the rod, the distance Δx between the points on the rod, and the length of time Δt between successive temperature measurements. Suppose that for $k = 0, 1, 2, \dots$, a vector $\mathbf{t}_k$ in $\mathbb{R}^4$ lists the temperatures at time $k\Delta t$. If the two ends of the rod are maintained at 0°, then the temperature vectors satisfy the equation $A\mathbf{t}_{k+1} = \mathbf{t}_k$ $(k = 0, 1, \dots)$, where

$$A = \begin{bmatrix} (1+2C) & -C & & \\ -C & (1+2C) & -C & \\ & -C & (1+2C) & -C \\ & & -C & (1+2C) \end{bmatrix}$$

a. Find the LU factorization of A when $C = 1$. A matrix such as A with three nonzero diagonals is called a *tridiagonal matrix*. The L and U factors are *bidiagonal matrices*.

b. Suppose $C = 1$ and $\mathbf{t}_0 = (10, 15, 15, 10)^T$. Use the LU factorization of A to find the temperature distributions $\mathbf{t}_1$, $\mathbf{t}_2$, $\mathbf{t}_3$, and $\mathbf{t}_4$.

[2] See Biswa N. Datta, *Numerical Linear Algebra and Applications* (Pacific Grove, CA: Brooks/Cole, 1994), pp. 200–201.

SOLUTION TO PRACTICE PROBLEM

$$A = \begin{bmatrix} 2 & -4 & -2 & 3 \\ 6 & -9 & -5 & 8 \\ 2 & -7 & -3 & 9 \\ 4 & -2 & -2 & -1 \\ -6 & 3 & 3 & 4 \end{bmatrix} \sim \begin{bmatrix} 2 & -4 & -2 & 3 \\ 0 & 3 & 1 & -1 \\ 0 & -3 & -1 & 6 \\ 0 & 6 & 2 & -7 \\ 0 & -9 & -3 & 13 \end{bmatrix}$$

$$\sim \begin{bmatrix} 2 & -4 & -2 & 3 \\ 0 & 3 & 1 & -1 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 10 \end{bmatrix} \sim \begin{bmatrix} 2 & -4 & -2 & 3 \\ 0 & 3 & 1 & -1 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = U$$

Divide the entries in each highlighted column by the pivot at the top. The resulting columns form the first three columns in the lower half of L. This suffices to make row reduction of L to I correspond to reduction of A to U. Use the last two columns of I_5

to make L unit lower triangular.

$$\begin{array}{ccc}
\begin{bmatrix}2\\6\\2\\4\\-6\end{bmatrix} & \begin{bmatrix}3\\-3\\6\\-9\end{bmatrix} & \begin{bmatrix}5\\-5\\10\end{bmatrix} \\
\div 2 & \div 3 & \div 5 \\
\downarrow & \downarrow & \downarrow
\end{array}$$

$$\begin{bmatrix}1 & & & \\ 3 & 1 & & \\ 1 & -1 & 1 & \cdots \\ 2 & 2 & -1 & \\ -3 & -3 & 2 & \end{bmatrix}, \quad L = \begin{bmatrix}1 & 0 & 0 & 0 & 0\\ 3 & 1 & 0 & 0 & 0\\ 1 & -1 & 1 & 0 & 0\\ 2 & 2 & -1 & 1 & 0\\ -3 & -3 & 2 & 0 & 1\end{bmatrix}$$

2.6 THE LEONTIEF INPUT–OUTPUT MODEL

WEB

Linear algebra played an essential role in the Nobel prize–winning work of Wassily Leontief, as mentioned at the beginning of Chapter 1. The economic model described in this section is the basis for more elaborate models used in many parts of the world.

Suppose a nation's economy is divided into n sectors that produce goods or services, and let $\mathbf{x}$ be a **production vector** in $\mathbb{R}^n$ that lists the output of each sector for one year. Also, suppose another part of the economy (called the *open sector*) does not produce goods or services but only consumes them, and let $\mathbf{d}$ be a **final demand vector** (or **bill of final demands)** that lists the values of the goods and services demanded from the various sectors by the nonproductive part of the economy. The vector $\mathbf{d}$ can represent consumer demand, government consumption, surplus production, exports, or other external demands.

As the various sectors produce goods to meet consumer demand, the producers themselves create additional **intermediate demand** for goods they need as inputs for their own production. The interrelations between the sectors are very complex, and the connection between the final demand and the production is unclear. Leontief asked if there is a production level $\mathbf{x}$ such that the amounts produced (or "supplied") will exactly balance the total demand for that production, so that

$$\left\{\begin{matrix}\text{amount}\\ \text{produced}\\ \mathbf{x}\end{matrix}\right\} = \left\{\begin{matrix}\text{intermediate}\\ \text{demand}\end{matrix}\right\} + \left\{\begin{matrix}\text{final}\\ \text{demand}\\ \mathbf{d}\end{matrix}\right\} \tag{1}$$

The basic assumption of Leontief's input–output model is that for each sector, there is a **unit consumption vector** in $\mathbb{R}^n$ that lists the inputs needed *per unit of output* of the sector. All input and output units are measured in millions of dollars, rather than in quantities such as tons or bushels. (Prices of goods and services are held constant.)

As a simple example, suppose the economy consists of three sectors—manufacturing, agriculture, and services—with unit consumption vectors $\mathbf{c}_1$, $\mathbf{c}_2$, and $\mathbf{c}_3$, as shown in the table that follows.